BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

This is a blank page.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 15/12, 15/09, 9/64, C07K 13/00, A61K 39/385, 39/39, 48/00

A1

(11) International Publication Number: WO 95/27058

(43) International Publication Date: 12 October 1995 (12.10.95)

(21) International Application Number: PCT/US94/03591

(22) International Filing Date: 1 April 1994 (01.04.94)

(71) Applicant *(for all designated States except US)*: THE UNIVERSITY OF UTAH [US/US]; Technology Transfer Office, Suite 170, 421 Wakara Way, Salt Lake City, UT 84108 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: RECHSTEINER, Martin, C. [US/US]; 714 North Medical Plaza, Salt Lake City, UT 84112 (US). REALINI, Claudio, A. [CH/US]; Apartment 1C, 637 South 600 East, Salt Lake City, UT 84102 (US).

(74) Agents: HOWARTH, Alan, J. et al.; Thorpe, North & Western, Suite 200, 9035 South 700 East, Sandy, UT 84070 (US).

(81) Designated States: AU, CA, CZ, FI, JP, NO, NZ, PL, RU, UA, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: MOLECULAR CLONING AND EXPRESSION OF A γ-INTERFERON INDUCIBLE ACTIVATOR OF THE PROTEASOME

(57) Abstract

Molecular cloning and expression of a human gene encoding a polypeptide activator of proteasomes is disclosed. The expressed activator has an M_r of about 29,000 and is functional in activating proteasomes *in vitro*. *In vivo* this activator polypeptide is inducible with γ-interferon in HeLa cells and occurs with a non-γ-interferon-inducible polypeptide with an M_r of about 31,000 in a hexameric activator compl x. The activator protein contains a lysine and glutamate rich region termed a KEKE motif. The KEKE motif appears to promote association between proteins and selection of peptides for presentation on MHC Class I receptors. A method for enhancing cell-mediated immunity against or tolerance to a selected immunogenic peptide is described comprising expressing activator and the selected peptide, wherein the selected peptide is adjacent to a KEKE motif, in an appropriate cell.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria
AU	Australia
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CS	Czechoslovakia
CZ	Czech Republic
DE	Germany
DK	Denmark
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgystan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakhstan
LI	Liechtenstein
LK	Sri Lanka
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SI	Slovenia
SK	Slovakia
SN	Senegal
TD	Chad
TG	Togo
TJ	Tajikistan
TT	Trinidad and Tobago
UA	Ukraine
US	United States of America
UZ	Uzbekistan
VN	Viet Nam

MOLECULAR CLONING AND EXPRESSION OF A γ-INTERFERON INDUCIBLE ACTIVATOR OF THE PROTEASOME

Background of the Invention

This invention was made with government support under Grant No. GM 37009 awarded by the National Institutes of Health. The government has certain rights in the invention.

This invention relates to multicatalytic proteases. More particularly, this invention relates to molecular cloning and expression in bacteria of a human gene encoding a γ-interferon-inducible activator of proteasomes. The activator protein contains a lysine and glutamate-rich region, termed a KEKE motif, that appears to promote association between KEKE-motif containing proteins and presentation of immunogenic peptides on MHC Class I receptors. The invention further relates to eliciting cellular immunity against or tolerance to selected immunogenic peptides.

Multimeric, ATP-dependent proteins serve important regulatory functions in both prokaryotic and eukaryotic cells. Two distinct *E. coli* proteases, Lon and Clp, have been shown to degrade specific regulatory proteins, thereby controlling a variety of bacterial processes. M. Maurizi, 48 Experientia 178 (1992). Only one ATP-dependent protease has been identified in nuclear or cytosolic extracts from eukaryotes. R. Hough et al., 261 J. Biol. Chem. 2400 (1986); R. Hough et al., 262 J. Biol. Chem. 8303 (1987). This large (26 S) protease degrades proteins conjugated to ubiquitin (Ub), M. Rechsteiner et al., 268 J. Biol. Chem. 6065 (1993), and is able to degrade unmodified ornithine decarboxylase complexed to antizyme, Y. Murakami et al., 360 Nature 597 (1992). Because of its involvement in Ub-mediated proteolysis, M. Rechsteiner, 66 Cell 615 (1991), the 26 S protease plays an important role in cell-cycle traverse, M. Glotzer et al., 349 Nature 132 (1991); A Hershko et al., 266 J. Biol. Chem. 16379 (1991), and gene

expression, M. Hochstrasser et al., 88 Proc. Nat'l Acad. Sci. USA 4606 (1991).

Confirming a model proposed by Hough et al., ATP/Ubiquitin-dependent Proteases, in Ubiquitin 101 (M. Rechsteiner ed., 1988), the 26 S protease is formed from a proteolytic core provided by the 20 S proteasome (also known as macropain, multicatalytic protease, or 20 S protease). E. Eytan et al., 86 Proc. Nat'l Acad. Sci. USA 7751 (1989); J. Driscoll & A. Goldberg, 265 J. Biol. Chem. 4789 (1990). Proteasomes are high molecular weight, multisubunit proteases that display a number of unusual structural and functional properties. These enzymes have been identified in every examined species from archaebacteria to humans. They have a native molecular weight of about 650,000 and a distinctive cylinder-shaped morphology in electron micrographs. A. Rivett, 268 Arch. Biochem Biophys. 1 (1989); M. Orlowski, 29 Biochemistry 10289 (1990). These cylinders measure 11 X 16 nm in outer dimensions with a central pore measuring about 2 nm in diameter, F. Kopp et al., 872 Biochim. Biophys. Acta 253 (1986), and comprise a stack of four rings, each ring containing six to eight low molecular weight subunits. Analysis of the subunits shows that most of them are electrophoretically distinct and range in molecular weight from about 20,000 to 35,000. S. Wilk & M. Orlowski, 40 J. Neurochem. 842 (1983); B. Dahlmann et al., 228 Biochem. J. 171 (1985). Individual proteasome subunits can be grouped into two families, termed α and ß, based on their similarity to the simpler archaebacterial enzyme. P. Zwickl et al., 31 Biochem. 964 (1992). Eukaryotic proteasome subunits have been shown by sequence analysis of cDNAs to represent the products of at least 13 different genes. K. Tanaka et al., 4 New Biologist 173 (1992). Surprisingly, the subunits are homologous to one

another, but not to any other known protease. Moreover, there is strong amino acid sequence similarity among subunits of various species. Proteasomes, therefore, may represent protein complexes composed of an evolutionarily related group of novel proteases.

Proteasomes also display catalytic features that are not closely similar to previously described proteases. For example, classification of proteasome active sites with protease inhibitors does not lend to easy assignment to one of the major protease families. A. Rivett, 268 Arch. Biochem Biophys. 1 (1989); M. Orlowski, 29 Biochemistry 10289 (1990); S. Wilk & M. Orlowski, 40 J. Neurochem. 842 (1983); B. Dahlmann et al., 228 Biochem. J. 171 (1985). Evidence suggests that proteasomes have three or more distinct proteolytic activities. S. Wilk & M. Orlowski, 40 J. Neurochem. 842 (1983); A. Rivett, 264 J. Biol. Chem. 12215 (1989); J. Arribas & J. Castaño, 265 J. Biol. Chem. 13969 (1990); B. Yu et al., 266 J. Biol. Chem. 17396 (1991). However, characterization of these activities in terms of catalytic mechanisms and subunit localization has not been achieved.

Proteasomes seem to play an obligatory role in the ubiquitin pathway of intracellular protein degradation. Hough et al., 261 J. Biol. Chem. 2400 (1986); R. Hough et al., 262 J. Biol. Chem. 8303 (1987). Yeasts with mutant proteasomes exhibit both a decreased rate of degradation of normal short-lived and abnormal proteins and an accumulation of ubiquitinated proteins. Proteasomes have also been implicated in ATP-dependent, ubiquitin-independent pathways of protein degradation and in antigen presentation on cell surfaces by major histocompatibility complex (MHC) glycoproteins, A. Townsend & H. Bodmer, 7 Ann. Rev. Immunol. 601 (1989); G. van Bleek & S. Nathenson, 2 Trends Cell Biol. 202

(1992); A. Goldberg & K. Rock, 357 Nature 375 (1992); J. Howard, 90 Proc. Nat'l Acad. Sci. USA 3777 (1993); J. Trowsdale, 9 Trends in Genetics 117 (1993), although their contribution to these various processes is unclear. Despite the probable importance of proteasomes in intracellular protein degradation, the mechanisms by which it mediates this function is unclear because several features of purified proteasomes differ significantly from features that characterize proteasome-mediated degradative pathways. Activation of proteasomes occurs during enzyme purifications in the absence of glycerol, and purified latent proteasomes can be activated directly *in vitro* by treatment with polycations, low concentrations of SDS, fatty acids, or dialysis against water.

These effects probably mimic some type of physiological activation, and at least three such activation mechanisms have been suggested. First, a specific type of latent 20 S proteasome appears to be activated directly by ATP. J. Driscoll & A. Goldberg, 86 Proc. Nat'l Acad. Sci. USA 787 (1989). Second, latent 20 S proteasomes can be activated by association with at least two poorly characterized proteins. In the presence of ATP, the cylindrical proteasome associates with an ATPase complex containing 15 or more different polypeptides to form the 26 S enzyme. L. Hoffman et al., 267 J. Biol. Chem. 22362 (1992). Assembly generates an enzyme capable of degrading Ub conjugates and results in elevated peptidase activity. Third, a simpler protein complex capable of stimulating the proteasome's peptidase activity has recently been described. M. Chu-Ping et al., 267 J. Biol. Chem. 10515 (1992); W. Dubiel et al., 267 J. Bio. Chem. 22369 (1992). A protein (PA28) that greatly stimulates the multiple peptidase activities of 20 S proteasomes has been purified from bovine red blood cells and bovine heart.

M. Chu-Ping et al., Identification, Purification, and Characterization of a Protein Activator (PA28) of the 20 S Proteasome (Macropain), 267 J. Biol. Chem. 10515 (1992). PA28 is a single polypeptide with an apparent subunit molecular weight of 28,000 as estimated by denaturing gel electrophoresis, and a native molecular weight of about 180,000 as estimated by gel filtration chromatography and density gradient centrifugation. Thus, the native activator may be a hexamer of a 28 kD polypeptide. PA28 apparently binds to proteasomes and may be a novel polypeptide because comparison of unpublished partial amino acid sequence data with the PIR, W. Barker et al., 20 Nucleic Acids, Res. 2023 (1992), and Swiss-Prot, A. Bairoch & B. Boeckmann, 20 Nucleic Acids Res. 2019 (1992), databases showed no significant similarities with any known protein. PA28 regulates three peptidase activities of proteasomes, including increasing the maximal reaction velocity and decreasing the half-maximal velocity. PA28 failed, however, to stimulate proteolysis of large protein substrates such as casein and lysozyme.

A protein complex has also been purified from human red blood cells that activates proteasomes. W. Dubiel et al., Purification of an 11 S Regulator of the Multicatalytic Protease, 267 J. Biol. Chem. 22369 (1992). The complex has an apparent molecular weight of about 200,000 on nondenaturing gels and consists of two protein species that migrate as a close doublet having M_rs of about 31,000 and 29,000 on denaturing electrophoretic gels. The two proteins are present in approximately equal concentrations, and proteasome-activating activity corresponds with the complex containing both of the subunits. The activator complex lacks intrinsic peptidase activity, but stimulates proteolysis of certain substrates about 60-fold, although activated proteasomes are unable to degrade ubiquitin-lysozyme conjugates, bovine serum

albumin, or lysozyme. Activation involves reversible binding of the activator complex to proteasomes.

Objects and Summary of the Invention

It is an object of the present invention to provide a molecular clone of the human gene for a proteasome activator.

It is also an object of the invention to provide a functional expressed protein derived from the cloned human gene for a proteasome activator.

It is another object of the invention to provide a method for activating proteasomes *in vitro*.

It is still another object of the invention to provide a method for producing selected amounts of immunogenic peptides for presentation on MHC I receptors.

It is yet another object of the invention to provide a method for inducing cell-mediated immunity against or tolerance to specific epitopes using plasmids encoding a proteasome activator and appropriate peptides for presentation adjacent to peptides that mark the peptides to be presented.

These and other objects are achieved by providing a purified polynucleotide having a nucleotide sequence that encodes a proteasome activator, wherein the activator is of human origin and has an M_r of about 29,000. The polynucleotide has a nucleotide sequence identified as SEQ ID NO:9 and encodes a protein having an amino acid sequence identified as SEQ ID NO:10. A protein capable of activating proteasomes *in vitro* is also provided, wherein the protein is expressed from the polynucleotide encoding the proteasome activator.

A method of activating proteasomes is also provided, the method comprising the step of contacting the proteasomes with the protein expressed from the polynucleotide encoding the proteasome activator under

conditions suitable for binding of the protein to the proteasomes.

A method for inducing synthesis in cultured human cells of an activator of proteasomes is further provided, wherein the activator comprises a hexameric activator complex including polypeptides having an M_r of about 29,000, the method comprising the step of treating the human cells with an effective amount of γ-interferon.

A method for enhancing cell-mediated immunity or tolerance to selected epitopes, such as epitopes from pathogens, is also provided. This method comprises co-expression of proteasome activator and appropriate precursors of presented peptides bearing the epitopes. The peptides to be presented are marked for presentation by adjacent lysine and glutamate rich peptides, termed KEKE motifs. In one illustrative embodiment, a plasmid is provided containing the gene for the 29 kD proteasome activator and an appropriate promoter and other signals for in vivo expression of the activator protein. A second plasmid is also provided containing nucleotide sequences encoding a KEKE motif adjacent to a peptide cassette and a carrier protein. The peptide cassette can contain immunogenic peptides selected from known pathogen proteins on the basis of their ability to bind MHC Class I receptors. The plasmids are injected into mammalian muscle according to known methods for producing cellular immunity to the selected pathogens. Production of cellular immunity or tolerance is selectable according to the amounts of epitope entering the presentation pathway.

Brief Description of the Drawings

FIGS. 1A and 1B show *in vivo* expression of the molecularly cloned human activator gene in *E. coli* by SDS-PAGE (FIG. 1A) and immunoblotting (FIG. 1B).

FIGS. 2A (nondenaturing gel assay), 2B (fluorometric assay), and 2C (2-dimensional PAGE analysis) show stimulation of peptide hydrolysis by recombinant human activator expressed in *E. coli*.

FIGS. 3A and 3B are, respectively, a Lineweaver-Burke plot of suc-Leu-Leu-Val-Tyr-MCA (SEQ ID NO:18) hydrolysis in the presence or absence of recombinant human activator and a plot of substrate-dependent stimulation of proteasome activity in the presence of recombinant human activator.

FIGS. 4A shows the position of activator from human red blood cells on a stained 2-dimensional gel. FIGS. 4B and 4C show, respectively, autoradiograms of HeLa cell proteins synthesized in the absence and presence of γ-interferon, showing the position of activator.

FIG. 5 shows amino acid sequences of "KEKE motifs" from human proteasome activator and certain proteasome subunits and chaperonins.

FIG. 6 is a diagramatic representation of a proteasome and its association with activator complexes to form an activated proteasome and with ATPase complexes to form a 26 S protease, each multisubunit structure having KEKE motif-containing peptides extending therefrom.

FIG. 7 shows release of MCA plotted as a function of time when proteasomes, activator complex, and a fluorogenic peptide substrate (SEQ ID NO:18) were mixed with either a ubiquitin-KEKE motif fusion peptide (SEQ ID NO:21) or ubiquitin.

FIG. 8 is a diagramatic representation of an activated proteasome in relation to other components of the antigen presentation pathway.

Detailed Description of the Invention

Before the present γ-interferon-inducible activator of proteasomes is disclosed and described,

it is to be understood that this invention is not limited to the particular process steps and materials disclosed herein as such process steps and materials may vary somewhat. It is also to be understood that the terminology used herein is used for the purpose of describing particular embodiments only and is not intended to be limiting since the scope of the present invention will be limited only by the appended claims and their equivalents.

In describing and claiming the present invention, the following terminology will be used in accordance with the definitions set out below.

As used herein, "activator complex" means the hexameric complex with a molecular weight of about 200,000 that can be purified from human cells and is capable of activating proteasomes *in vitro*. Activated proteasomes are to be contrasted with latent proteasomes, which are incapable of proteolytic activity. Thus, activated proteasomes are capable of proteolytic activity through the mediation of an activator. Activator complexes contain two subunits having M_rs of about 31,000 and about 29,000, which are referred to herein as the 31 kD and 29 kD subunits, respectively.

As used herein, "recombinant 29 kD activator," "recombinant activator," and similar terms mean the protein produced by molecular cloning of the human gene for the 29 kD subunit of the activator complex, transfer of the cloned gene to a cell system for expression of foreign proteins, and expression of the cloned gene in the cell system.

As used herein, "transformable polynucleotide" means a plasmid, phagemid, cosmid, viral nucleic acid, and the like that can be transferred, transformed, or transfected into host cells and be physiologically active therein.

As used herein, by an "effective amount" of γ-IFN is meant the amount of γ-IFN necessary to elicit the selected induction of the 29 kD subunit of the activator complex.

As used herein, "PCR" means polymerase chain reaction, the process for *in vitro* amplification of DNA disclosed in U.S. Patent Nos. 4,683,195 and 4,683,202.

Molecular Cloning of 29 kD Activator Subunit

Activator complex was partially purified from human red blood cells as described previously, W. Dubiel et al., Purification of an 11 S Regulator of the Multicatalytic Protease, 267 J. Biol. Chem. 22369 (1992), and then was gel purified by SDS-PAGE, 3 J. Sambrook et al., Molecular Cloning: A Laboratory Manual § 18.47 (2d ed., 1989). The purified activator complex proteins were then subjected to cleavage with V8 protease or with CNBr according to standard procedures. The resulting peptides were fractionated by HPLC and sequenced using an ABI automated gas-phase sequencer. W. Dubiel et al., 267 J. Biol. Chem. 22699 (1992). Partial amino acid sequences were thus obtained for both the 31 kD and the 29 kD subunits of the activator complex. SEQ ID NO:1 through SEQ ID NO:5 disclose partial amino acid sequences from cleavage products of the 29 kD subunit of the activator complex. These peptide sequences were used to design sense and anti-sense degenerate oligonucleotide PCR primers having the sequences identified, respectively, as SEQ ID NO:6 and SEQ ID NO:7, where the nucleotides identified as N were inosine residues. These oligonucleotides were combined in PCR reactions with $CsCl_2$-purified total RNA from HeLa cells, and PCR was performed using "GENE AMP" components (Perkin Elmer Cetus). Amplified DNA was separated on agarose gels and the appropriate

products were identified upon hybridization with the [γ-^{32}P]-labeled oligonucleotides of SEQ ID NO:8 (where the nucleotide identified as N was inosine), located between the two primers. DNA that hybridized to SEQ ID NO:8 was subcloned into the EcoRI and HindIII sites of the plasmid sold under the trademark "pBluescript KS" (Stratagene, La Jolla, CA) and sequenced using the "SEQUENASE" kit (U.S. Biochemicals, Cleveland, OH). The sequences thus obtained were used to design non-degenerate oligonucleotides for screening cDNA libraries. Approximately 10^5 phage recombinants from a λgt11 cDNA library from human tonsils were screened with a non-degenerate hybridization probe (SEQ ID NO:86) produced by oligonucleotide synthesis and end-labeled with [γ-^{32}P]-ATP using T4-polynucleotide kinase (Boehringer Mannheim). The linear "λ ZAP II" (Stratagene) vector of positive recombinant bacteriophages was excised and recircularized into "pBluescript" phagemid according to the *in vivo* excision protocol of Stratagene. Inserts of positive λgt11 clones were subcloned into EcoRI sites of pBluescript phagemids. Both constructs were amplified in the XL Blue 1 strain (Stratagene) of *E. coli* and processed for DNA sequencing. The PC gene algorithm and database were used to analyze the nucleotide sequence (SEQ ID NO:9) and deduced amino acid sequence (SEQ ID NO:10).

The longest clone obtained through this screening procedure contained an open reading frame (ORF) for a 249 residue polypeptide (SEQ ID NO:10) with a calculated molecular weight of 27,330 daltons, in reasonable agreement with an apparent molecular weight of 29 kD for the smaller subunit of the activator complex. Each of the five sequenced peptides (SEQ ID NO:1 through SEQ ID NO:5) is present in SEQ ID NO:10.

A search of the PIR library, W. Barker et al., 20 Nucleic Acids Res. 2023 (1992), using the entire

sequence (SEQ ID NO:10) revealed an exact match to a recently submitted γ-interferon (γ-IFN) induced protein of unknown function. And as shown below, synthesis of the 29 kD subunit of the activator complex is increased 5-fold by γ-IFN. Since extensive sequence similarity with other known sequences was not found, the activator appears to be a novel protein.

Expression of Cloned 29 kD Activator Subunit

To determine if the 29 kD activator cDNA was full length and to initiate biochemical studies on the protein, the longest cDNA was subcloned into a pAED4 expression system (a gift from Dr. Tom Albers), wherein a cloned gene is expressed under control of the *lac* promoter. This cloning step was accomplished by ligation of the cDNA containing the gene for the 29 kD subunit of the activator complex into the NdeI and BamHI sites of the T7 polymerase-dependent expression vector pAED4. Ligation products were transformed into BL21(DE3) cells prepared for $CaCl_2$-dependent transformation. K. Shigekawa & W. Dower, 6 BioTechniques 742 (1988). Soluble protein fractions were obtained by sonication and centrifugation at 39,000 *g* for 30 min at 4°C. Recombinant *E. coli* were either induced for 2 hours with 0.5 mM isopropyl-ß-thiogalactopyranoside (IPTG), a gratuitous inducer of the *lac* promoter, or grown in the absence of IPTG prior to sonication. Proteins were subjected to electrophoresis on a 12% SDS-polyacrylamide gel and stained with Coomassie Brilliant Blue R (FIG. 1A) or transferred to a nitrocellulose membrane for immunoblot analysis (FIG. 1B). The nitrocellulose membrane was blocked for 60 min in 5% dried milk in TBS (25 mM Tris·HCl, pH 7.5, 0.9% NaCl, and 0.02% sodium azide). The filter was then incubated with mouse anti-human red cell activator complex serum (1/2000 dilution) for 12 hours at 4°C, washed in TBS,

incubated for 6 hours in the presence of [I^{125}]-rabbit anti-mouse IgG (DAKO), extensively washed in TBS, and exposed for 3 days to X-omat AR film (Kodak) at -20°C. Polyclonal antibodies were raised in Balb/C mice injected intraperitoneally with human red cell activator complex purified as described, W. Dubiel et al., Purification of an 11 S Regulator of the Multicatalytic Protease, 267 J. Biol. Chem. 22369 (1992), and hereby incorporated by reference.

FIG. 1A shows that high levels of a 29 kD protein were produced upon treatment of the recombinant *E. coli* with IPTG, and the induced protein comigrated with the 29 kD polypeptide component of human red cell activator complex on SDS-PAGE. The lane labeled "Human" contained partially purified human red blood cell activator. "Mix" contained recombinant 29 kD activator and partially purified human red blood cell activator that were mixed prior to electrophoresis. The lanes marked "+" and "-" contained, respectively, the soluble protein fractions from recombinant *E. coli* strain BL21(DE3) cells induced and not induced with IPTG. The lane marked "STD" contained molecular weight markers with molecular weights as indicated. FIG. 1B is an immunoblot of the gel from FIG. 1A and demonstrates that the induced protein reacts with antibodies against human red cell activator complex. Moreover, the recombinant 29 kD protein resolved on two-dimensional electrophoresis (not shown) as three species with pIs between 5.1 and 5.6, in excellent agreement with similar analyses on purified human red cell activator complex. W. Dubiel et al., Purification of an 11 S Regulator of the Multicatalytic Protease, 267 J. Biol. Chem. 22369 (1992). Thus, the recombinant 29 kD protein expressed in bacterial cells appears to be a faithful copy of the 29 kD subunit of the activator complex.

The recombinant protein formed inclusion bodies and was not functional in *E. coli* cells grown to high density and induced with 1 mM IPTG. However, fully soluble recombinant 29 kD activator was obtained by short periods of induction using lower levels of IPTG. FIG. 2A shows the results on peptide hydrolysis of mixing increasing amounts of *E. coli* extract induced at low IPTG concentrations (+ IPTG), *E. coli* extract from non-induced cells (- IPTG), or purified human red cell activator complex (human) with proteasomes. The mixtures were incubated for 10 min at 37°C with 300 ng of purified human proteasomes. The mixtures were then subjected to electrophoresis on 4.5% non-denaturing gels for 4.5 hours as described in L. Hoffman et al., 267 J. Biol. Chem. 22362 (1992), and hereby incorporated by reference. The gel was overlaid with 200 μM suc-Leu-Leu-Val-Tyr-MCA (succinyl-Leu-Leu-Val-Tyr-7-amido-4-methylcoumarin; SEQ ID NO:18), a fluorogenic synthetic peptide substrate, incubated for 10 min at 37°C, and the released 7-amido-4-methylcoumarin (MCA) was localized by UV transillumination. Recombinant 29 kD activator was prepared as described above in the discussion of FIGS. 1A and 1B. Proteasomes were prepared from outdated human blood as described in W. Dubiel et al., Purification of an 11 S Regulator of the Multicatalytic Protease, 267 J. Biol. Chem. 22369 (1992). The results show that addition of recombinant 29 kD activator in *E. coli* extracts to proteasomes led to progressive activation of cleavage of suc-Leu-Leu-Val-Tyr-MCA (SEQ ID NO:18) upon peptide overlay of native gels (FIG. 2A). FIG. 2B shows the results of a fluorometric assay of peptide hydrolysis in the presence of recombinant 29 kD activator. Various amounts of *E. coli* extract (0 μl, ■; 2 μl, ▲; 5 μl, ▼; 10 μl, □; 20 μl,) containing recombinant 29 kD activator were added to 400 ng of purified rabbit

proteasomes (purified as described above for human proteasomes), and the samples were incubated at 37°C in 100 μM of suc-Leu-Leu-Val-Tyr-MCA (SEQ ID NO:18). At the indicated times, 100 μl aliquots of the mixture were quenched with 200 μl of cold 100% ethanol and the fluorescence was measured by excitation at 380 nm and emission at 440 nm. The sample designated 0 μl contained proteasomes plus 20 μl of extract from uninduced cells. Fluorescence is plotted as a function of incubation time in the presence of increasing amounts of extract from IPTG-induced cells. The inset shows the stimulation of suc-Leu-Leu-Val-Tyr-MCA (SEQ ID NO:18) hydrolysis as a function of added recombinant activator for 5 min incubation at 37°C. Stimulation is defined as $S = (F_{Prot} + ACT)/F_{Prot}$, where F_{Prot} + ACT is the rate of change in the fluorescence F in the presence of a given amount of activator, and F_{Prot} is the measured fluorescence in the absence of activator. These results revealed about 25-fold stimulation by saturating amounts of *E. coli* extract containing recombinant activator (FIG. 2B).

FIG. 2C shows a 2-dimensional PAGE analysis of activator/proteasome association. Purified human proteasomes (400 ng) and partially purified recombinant 29 kD activator (10 μl) were mixed and subjected to electrophoresis for 6 hours at 4°C on an 8% non-denaturing polyacrylamide gel in TBE (90 mM Tris, 1.6 mM boric acid, 0.08 mM EDTA, pH 8.3) at a constant voltage of 10 V/cm. After electrophoresis, an individual lane from the gel was incubated for 10 min in 30 mM of Tris·HCl, pH 6.8, 1% SDS, 5% glycerol, and 5 mM 2-mercaptoethanol, and then loaded on a 10% SDS-polyacrylamide gel, U. Laemmli, 227 Nature 680 (1970). Proteins were stained with Coomassie Brilliant Blue R 250. The small vertical arrows indicate the relative position of proteins whose migration was not affected by proteasomes. The dotted

line compares the migration of recombinant 29 kD activator in the absence or presence of proteasomes, upper v. lower panels, respectively. These results show formation of a stable complex between recombinant 29 kD activator and proteasomes. Activation of peptide hydrolysis is known to result from binding of red cell activator complexes to proteasomes rather than modification of either component. W. Dubiel et al., Purification of an 11 S Regulator of the Multicatalytic Protease, 267 J. Biol. Chem. 22369 (1992). Thus, electrophoretic (FIGS. 2A and 2C) and fluorometric assays (FIG. 2B) show that the soluble recombinant 29 kD activator binds proteasomes and stimulates their peptidase activity.

Kinetic analyses provided further comparison between recombinant 29 kD activator and the molecule purified from human red cells. FIG. 3A shows a Lineweaver-Burke plot of suc-Leu-Leu-Val-Tyr-MCA (SEQ ID NO:18) hydrolysis in the absence of added activator (●) and presence of activator complex purified from human red cells (■) or recombinant 29 kD activator (○). Purified human proteasomes (400 ng) and 30 μl of recombinant 29 kD activator or partially purified human red cell activator complexes were incubated with varying concentrations of suc-Leu-Leu-Val-Tyr-MCA (SEQ ID NO:18) as indicated in the figure. Peptide hydrolysis was monitored with a Perkin-Elmer LS-5 Fluorescence Spectrophotometer with an excitation wavelength of 380 nm and an emission wavelength of 440 nm, and initial velocities (nmol/min/mg of proteasomes) were determined for each concentration of substrate. The double reciprocal plots in FIG. 3A reveal that both the recombinant 29 kD activator and the red cell activator complex increase V_{max} by 50-fold at 200 μM of suc-Leu-Leu-Val-Tyr-MCA (SEQ ID NO:18) and decrease the K_m for hydrolysis of this peptide from 60 μM to ~4 μM.

FIG. 3B shows that the stimulation of proteasome activity by activator protein is substrate dependent. Increasing amounts of recombinant 29 kD activator were added to 400 ng of purified human red cell proteasomes, and the mixture was incubated at 37°C with 100 μM of suc-Arg-Pro-Phe-His-Leu-Leu-Val-Tyr-MCA (□, SEQ ID NO:19), suc-Gly-Pro-Leu-Gly-Pro-MCA (■, SEQ ID NO:20), suc-Leu-Leu-Val-Tyr-MCA (●, SEQ ID NO:18), Cbz-Leu-Leu-Glu-pNA (○, benzyloxycarbonyl-Leu-Leu-Glu-*p*-nitroaniline), or Pro-Phe-Arg-MCA (▲). The reactions were quenched after 10 min, and the fluorescence was measured at 440 nm for MCA-containing substrates, or with excitation at 335 nm and emission at 410 nm for Cbz-Leu-Leu-Glu-pNA. Stimulation was determined as defined in the discussion of FIG. 2B. Recombinant activator stimulates hydrolysis of suc-Leu-Leu-Val-Tyr-MCA (SEQ ID NO:18) more than LLE-pNA (Leu-Leu-Glu-*p*-nitroaniline), and cleavage of these peptides is enhanced to a greater extent than cleavage of PFR-MCA (Pro-Phe-Arg-7-amido-4-methylcoumarin) (FIG 3B). This pattern is identical to that previously demonstrated for red cell activator complex. W. Dubiel et al., Purification of an 11 S Regulator of the Multicatalytic Protease, 267 J. Biol. Chem. 22369 (1992). Incubation of ^{35}S-methionine labeled recombinant 29 kD activator with proteasomes followed by SDS-PAGE and autoradiography gave no evidence for cleavage of the recombinant 29 kD activator, thereby eliminating the possibility that the recombinant 29 kD activator serves as a proteasome substrate. These physical and enzymatic tests show that recombinant activator is very similar in activity to the activator complex obtained from human red cells, and physically resembles the 29 kD subunit of activator complexes.

Induction with γ-Interferon

The deduced amino acid sequence of the 29 kD subunit of the activator complex matches exactly the

sequence of a γ-IFN-induced protein, and proteasomes have previously been implicated in antigen presentation. A. Townsend & H. Bodmer, 7 Ann. Rev. Immunol. 601 (1989); G. van Bleek & S. Nathenson, 2 Trends Cell Biol. 202 (1992); A. Goldberg & K. Rock, 357 Nature 375 (1992); J. Howard, 90 Proc. Nat'l Acad. Sci. USA 3777 (1993); J. Trowsdale, 9 Trends in Genetics 117 (1993). For these reasons, the effects of γ-IFN on activator synthesis were examined. Cells of the human HeLa line, D98/AH2, were plated at 2 X 10^6 per 25 cm^2 T-flask in McCoys medium containing 200 μg/ml recombinant γ-IFN (Chemicon) or no γ-IFN. After 72 hours, the cells were rinsed with F12 medium and further cultured in F12 medium lacking methionine, but containing 50 μCi/ml ^{35}S-methionine (700 Ci/mmole). Three hours later, the cells were recultured in McCoys medium for 1 hour prior to harvest by trypsinization. The cells were dissolved in focusing buffer and 2-dimensional PAGE was performed as described. W. Dubiel et al., Purification of an 11 S Regulator of the Multicatalytic Protease, 267 J. Biol. Chem. 22369 (1992). After fixation, the gels were dried onto Whatman 3M paper, which accounts for the mottled background in FIG. 4A. The dried gels were exposed to a Kodak XAR film for 6 days. FIG. 4A shows a Coomassie stain of a two-dimensional gel with the position of human red cell activator indicated by the arrowhead. FIGS. 4B and 4C are, respectively, autoradiograms showing two-dimensional separation of proteins synthesized in the absence and presence of γ-IFN. It is evident from visual examination of the autoradiograms in FIGS. 4B-C that synthesis of the three 29 kD species comprising the activator is markedly stimulated by γ-IFN. Phosphorimager analysis revealed that in cells exposed to γ-IFN, 5.7-fold more ^{35}S-methionine was incorporated into the activator.

Incorporation ratios (+/- γ-IFN) for seven reference proteins were 0.95, 2.0, 1.2, 1.0, 0.5, 0.9, and 1.1, indicating that the two gels were equally loaded. Thus, in addition to subunits encoded by LMP2 and LMP7, Y. Yang et al., 89 Proc. Nat'l Acad. Sci. USA 4928 (1992), the 29 kD subunit of the activator complex represents another γ-IFN-inducible component of proteasomes.

The amount of γ-IFN necessary to induce synthesis of the 29 kD subunit of the activator complex has not been determined. Further, it is possible that different cell types may exhibit different susceptibilities to γ-IFN. Therefore, by an "effective amount" of γ-IFN is meant the amount of γ-IFN necessary to elicit the selected induction of the 29 kD subunit of the activator complex.

The immunological, physical, and enzymatic tests presented herein provide considerable evidence that the proteasome activator protein cloned and expressed as cDNA is equivalent to activator complex purified directly from red blood cells. There is, however, a significant difference between the two preparations. Activator complex from red cells migrates on SDS-PAGE as a close doublet of 31 kD and 29 kD subunit proteins. Bovine red cell activator is a single 28 kD protein, M. Chu-Ping et al., 267 J. Biol. Chem. 10515 (1992), and a single 30 kD activator has been found in rabbit reticulocytes. These results raise the questions of whether the 31 kD and 29 kD subunit proteins in humans are modifications of a single polypeptide or are distinct proteins. The following observations demonstrate the existence of distinct proteins. First, the five peptides (SEQ ID NO:1 through SEQ ID NO:5) obtained from the 29 kD subunit protein are distinct from a partial amino acid sequence obtained from the 31 kD subunit protein.

Second, a partial cDNA has been obtained, presumably encoding the 31 kD subunit protein, that contains an ORF encoding a 235 residue protein that exhibits 48% amino acid sequence similarity to that of SEQ ID NO:10. Thus, it is virtually certain that human cells express two distinct proteins that comprise the activator complex. The molecule identified as SEQ ID NO:10 is the smaller of the two proteins in the SDS-PAGE doublet. By itself, it is capable of activating proteasomes, and it is induced by γ-IFN treatment of HeLa cells (FIGS. 4A-C). The 31 kD subunit of the activator complex is not induced by γ-IFN treatment of HeLa cells and, at present, it is not known whether this larger protein activates peptide hydrolysis by proteasomes.

Unique Sequence Motifs That Promote Protein Associations

An unusual and striking feature of the activator sequence (SEQ ID NO:10) is the lysine-glutamate rich region extending from lysine 70 to lysine 97 (SEQ ID NO:11). This "KEKE motif" or "KEKE sequence," named from the one letter code names for lysine (K) and glutamate (E), is particularly interesting because similar stretches of "alternating" glutamate and lysine residues, though rare among known proteins, are present in proteasome subunits C9 and 28.1, subunit 12 of the 26 S protease, as well as certain chaperonins (FIG. 5). KEKE sequences from the 29 kD activator (SEQ ID NO:11), proteasome subunit C9 (SEQ ID NO:12), and 26 S protease subunit S12 (SEQ ID NO:13) were used to search the PIR library for related sequences. To the sequences most similar to each query sequence, the following criteria were applied to determine whether they conformed to our definition of a KEKE motif: (1)

the sequence was 13 amino acids or longer; (2) greater than 60% of the residues were lysine, glutamate, or aspartate; (3) no more than four consecutive negatively-charged or positively-charged residues were present; and (4) the sequence was devoid of tryptophan, tyrosine, phenylalanine, and proline. Of the 100,346 entries in the PIR library (release 39), only 106 proteins fulfilled these criteria. Similar sequences composed of arginine (a basic amino acid that can often be substituted for lysine) and aspartate (an acidic amino acid usually considered equivalent to glutamate) are present in only two proteins, hnRNP70 and human RD protein. M. Levi-Strauss et al., 240 Science 201 (1988); R. Spritz et al., 15 Nucleic Acids Res. 10373 (1987). Thus, KEKE motifs are not simply statistically expected arrangements of amino acids.

It is well documented that proteasomes associate with other proteins, including activator complex and a regulatory ATPase complex (AC). L. Hoffman et al., 267 J. Biol. Chem. 22362 (1992); M. Chu-Ping et al., 267 J. Biol. Chem. 10515 (1992); W. Dubiel et al., 267 J. Biol. Chem. 22369 (1992); A. Udvardy, 268 J. Biol. Chem. 9055 (1993). FIG. 6 illustrates these interactions, wherein is shown a proteasome 10 with KEKE motif-containing peptides 12 extending therefrom. Hexameric activator complex 14, also having KEKE motif-containing peptides 12 extending therefrom, binds to the proteasome 10 to form an activated proteasome 16. FIG. 6 also shows an AC complex 18, with KEKE motif-containing peptides 12 extending therefrom, binding to the proteasome 10 to form a 26 S protease 20. Each of these multisubunit complexes contains at least one component with a strong KEKE motif (FIG. 5). Under certain circumstances,

proteasomes form a stable complex with Hsp 90, which also exhibits a KEKE motif (FIG. 5). Thus, four proteins containing KEKE sequences bind one another. Although such interactions may not involve the KEKE motifs per se, these results suggest that KEKE motifs are responsible for the observed associations.

The following experiment provides additional evidence for KEKE motif-mediated association of proteins. A ubiquitin-KEKE motif fusion peptide (SEQ ID NO:21) was prepared, Y. Yoo et al., 264 J. Biol. Chem. 17078 (1989), using the KEKE motif sequence from the 29 kD subunit of activator complex (SEQ ID NO:11). This fusion peptide (SEQ ID NO:21) and ubiquitin were tested separately for binding to red cell activator complex prepared as described above. Constant amounts of proteasomes and activator complex were mixed with either the ubiquitin-KEKE motif fusion peptide (SEQ ID NO:21) or ubiquitin. Degradation of the fluorogenic peptide suc-Leu-Leu-Val-Tyr (SEQ ID NO:18) was monitored with a Perkin-Elmer LS-5 fluorescence spectrophotometer using excitation at 380 nm and emission at 44 nm. FIG. 7 shows release of MCA plotted against incubation time, wherein MCA is released at a relatively constant rate from the fluorogenic peptide (SEQ ID NO:18) in the presence of the ubiquitin-KEKE motif fusion peptide (SEQ ID NO:21), but little or no MCA is released in the presence of ubiquitin. Thus, addition of a KEKE motif to a polypeptide not otherwise able to bind to a KEKE motif-containing protein conferred the ability to bind to that KEKE motif-containing protein.

Role of KEKE Motif in Antigen Presentation

Current views of antigen presentation by Class I receptors encoded in the major histocompatibility

locus (MHC) invoke cytosolic proteolysis of cellular, viral, or other parasitic proteins to produce peptides for presentation on cell surfaces. J. Yewdell & J. Bennink, 52 Adv. in Immunol. 1 (1992); Rammensee et al., 11 Ann. Rev. Immunol. 213 (1993); R. Germain, 76 Cell 287 (1994). These peptides are transported into the lumen of the endoplasmic reticulum (ER), where they bind a groove in the MHC I receptor that accomodates protein fragments of 8 to 10 amino acids in length. Association of the MHC I receptor with a tight-binding peptide and $ß_2$-microglobulin releases the receptor from calnexin, an 88 kD chaperonin embedded in the ER membrane. The MHC I:peptide:$ß_2$-microglobulin complex is then transported to the cell surface. A number of questions concerning this presentation pathway are unresolved and somewhat controversial, such as which protease(s) generates the peptides, how large are the peptides, how are the peptides selected for presentation, and so forth.

The results presented herein focus on how peptides are selected from precursor proteins for presentation by MHC I receptors. Cells typically display about 50,000 MHC I receptors on their surfaces. Most peptides bound to MHC I receptors are present at between about 10 to about 1000 copies per cell. There are, however, about 10,000 copies of the peptide Ser-Phe-Phe-Pro-Glu-Ile-Thr-His-Ile (SEQ ID NO:22) bound to MHC I receptors on p815 cells. H. Rammensee et al., 11 Ann. Rev. Immunol. 213 (1993). SEQ ID NO:22, which originates from JAK 1 kinase, is found in the sequence identified as SEQ ID NO:23, which contains a strong KEKE motif adjacent to the presented peptide, SEQ ID NO:22. Given the proximity of such a highly presented peptide to a strong KEKE motif, the sequence context of a large number of

presented peptides was examined. Presented peptide sequences were obtained from K. Falk & O. Rotzschke, 5 Immunol. 81 (1993); H. Rammensee et al., 11 Ann. Rev. Immunol. 213 (1993); T. Jardetsky, 353 Nature 326 (1991); D. Hunt et al., 255 Science 1261 (1992); M. Di Brino et al., 152 J. Immunol. 620 (1994); and M. Corr et al., 176 J. Exper. Med. 1681 (1992).

Results of this survey of 51 presented peptides are presented in Tables 1-3. Twelve of the presented peptides examined originate from proteins that contain a KEKE sequence motif as defined above (Table 1). Another 6 of the presented peptides contain a KEKE-like sequence (Table 2), wherein a KEKE-like sequence is defined as conforming to the definition of a KEKE motif except that the proportion of lysines and glutamates is slightly less than 60% and/or there is one excluded amino acid residue present. Presented peptides not associated with KEKE motifs or KEKE-like sequences are shown in Table 3. Inasmuch as the abundance of KEKE motifs in the PIR library is only about 0.1% (106/100,346), there is at least a 150-fold enrichment for KEKE motifs, strictly defined, in proteins that generate MHC I presented peptides.

Table 1	
Presented Peptides Associated with KEKE Motifs	
Source	SEQ ID NO:
JAK1	23
HSP90	26
BBC1	30
eEF2	29
Spectrin	31
Plasmodium yoelii	35, 36
Plasmodium berghei	37, 38
Plasmodium knowlsei	39, 40
HIV gag	43, 44
IL6 precursor	54, 55
HSP90	64, 65
BBC1	66

Table 2	
Presented Peptides Associated with KEKE-like Motifs	
Source	SEQ ID NO:
BIP	47, 48
PGK	51, 52
Restin	57, 58, 59
Plasmodium falciparum	41, 42
Polyoma virus T antigen	87, 88
T.A. P198	67

Table 3

Presented Peptides Not Associated with KEKE Sequences

Source	SEQ ID NO:	Source	SEQ ID NO:
Ovalbumin	63	P91A	28
Fibrillarin	56	PNAS Dec. 93	27
Sendai virus	68	Ornithine decarboxylase	33
Influenza matrix	60	VSV nuclear protein	34
Ribosome S16	69	p68	45
HTLV-1	70	Listeriol.	46
Influenza hemaglutinin	71	Herpes virus	50
TIS21	72	Ribosome L18	53
MAG-1 antigen	73	Polyoma T antigen	61
Influenza hemaglutinin	74	Polyoma T antigen	62
Measles fusion protein	75	Influenza virus	85
Tristetraproline	76	HIV Gp160	84
PPAS (yeast)	77	Ribosome L28	83
Influenza hemagglutinin	78	Ribosome L8	82
E1A 32 kD protein	79	HLA CW-3	81
Influenza (A, JAP)	80	Influenza hemagglutinin	32
LMCV nuclear protein	49		

Self-association of KEKE motifs is thought to account for their significant enrichment in components of the 26 S protease/proteasome proteolytic pathway and for their enrichment in proteins known to bear presented peptides (see FIG. 8). Two of the six or more α subunits of the proteasome contain C-terminal KEKE motif sequence extensions 12. Each subunit of the activator complex 14 possesses a KEKE motif. Thus, it is thought that activator complex 14 binds a proteasome 10 with KEKE motif sequences to spare. These excess KEKE sequences in the activator complex 14 are thought to be available to bind KEKE or KEKE-like regions in potential proteolytic substrates. The presence of a 40 residue KEKE sequence (SEQ ID NO:24) in the cytoplasmic tail of calnexin 22 is also thought to play a role in this pathway. As mentioned above, this chaperonin holds MHC I receptors 24 in the membrane 26 of the endoplasmic reticulum until immunogenic peptides bind the MHC I receptors 24. Thus, the presence of six KEKE sequences in the hexameric activator complex 14 provides for its simultaneous association with a proteasome 10, a protein substrate, and a calnexin molecule 22. Binding of the protein substrate and proteasome 10 facilitates cleavage of the protein substrate by the proteasome 10. After this, the activator complex:proteasome:peptide complexes engage calnexin 22, and the peptides are released for transfer to the lumen 28 of the ER, presumably by TAPs (transporters associated with antigen presentation) 30. Thus, the activator complex 14 channels available proteasomes 10 into the antigen presentation pathway. This is consistent with the data described above wherein γ-IFN induces synthesis of the activator complex in HeLa cells, as has been observed for other components of the antigen presentation pathway, e.g. TAP1/TAP2, MHC-

Class I and Class II molecules, and LMP2 and LMP7 subunits of the proteasome.

The structural basis for interactions among KEKE sequences is not yet understood. Nevertheless, a similar sequence motif in caldesmon (SEQ ID NO:17) is likely to be a helix. Accordingly, it is thought that KEKE sequences may form helical bundles. Computer analysis of the KEKE motif predicts this region to form a very hydrophilic α-helix. Proline residues, which destabilize α-helices, are absent from the activator KEKE motif (SEQ ID NO:11), but enriched in both flanking regions, e.g. prolines 60, 64, 66, and 68 and occupy the N-terminal edge and prolines 99, 100, and 103 are present at the C-terminal boundary.

Method for Enhancing Cell-Mediated Immunity

The foregoing disclosure on how the activator and its KEKE regions promote presentation of peptides by MHC Class I receptors suggests a method for inducing high levels of cell-mediated immunity against or tolerance to specific pathogen-encoded peptides in warm-blooded animals including humans. The procedure requires vigorous production of the proteasome activator in the cytosol of antigen presenting cells. At the same time, the cell must be synthesizing reasonable amounts of the immunologic peptide in a precursor that possesses one or more adjacent KEKE motifs. The surface abundance of immunologic peptides affects whether immunity or tolerance to the peptides is elicited. P. Allen, Peptides in Positive and Negative Selection: A Delicate Balance, 76 Cell 593 (1994). For example, medium surface abundance of an immunologic peptide can trigger positive selection of specific T cells and, hence, immunity to the peptide, whereas high surface abundance of the peptide can trigger negative selection of specific T cells resulting in tolerance to the peptide. Synthesis of

peptides to elicit immunity or tolerance can be achieved, illustratively, by producing two eukaryotic expression plasmids that encode the activator and the precursor, although the invention lies in co-expression of activator, the presented peptide, and a KEKE motif positioned adjacent to the presented peptide, and not in the specific examples that follow nor the specific plasmids used.

The cDNA for activator can be cloned into unique cloning sites of the eukaryotic expression vector pSG5 (Stratagene). Cloning sites can be generated by PCR, and the PCR products directly ligated into the expression vector using standard recombinant techniques. The plasmid pSG5 contains the early SV40 promoter, ß-globin intron II, and a signal for poly(A) tail production to improve the level of <u>in vivo</u> expression of the inserted protein gene. Thus, the resulting expression plasmid, termed plasmid I, will induce synthesis of large amounts of activator upon introduction into suitable host cells.

Oligonucleotides encoding the candidate immunogenic peptides would be cloned into a second eukaryotic expression vector, also using the pSG5 expression vector. The candidate immunogenic peptides are selected from known pathogen proteins on the basis of their ability to bind Class I receptors. Examples of such immunogenic peptides include influenza hemagglutinin (SEQ ID NO:32) and matrix (SEQ ID NO:60) proteins, VSV nuclear protein (SEQ ID NO:34), *Plasmodium falciparum* protein (SEQ ID NO:41), and HIV gag protein (SEQ ID NO:44). The protein expressed by this plasmid, termed plasmid II, has the following structure:

Met-Ala-Ala-(KEKE motif sequence)-(peptide cassette)-Ala-Ala-(carrier protein). The KEKE motif sequence could be any KEKE motif now known or later identified, but SEQ ID NO:11, the KEKE motif from activator is

preferred. It is possible that KEKE motifs have different strengths for enhanced antigen presentation, or other types of specificities not now recognized, thus fine-tuning of antigen presentation may be possible through selection of KEKE motifs used in the plasmids. For example, selection of appropriate KEKE motifs can be used for selectively inducing tolerance or immunity to the presented peptide based on the amount of peptide that enters the presentation pathway. The peptide cassette comprises Glu-Glu-Val followed by 8-10 amino acids of any specified immunogenic peptide, such as mentioned above. The carrier protein element of the construct is added to increase the size of the expressed chimeric protein. Many smaller peptides are rapidly degraded inside cells, thus it thought that residence time in the cell can be increased by fusion to a carrier protein. Dihydrofolate reductase (DHFR) is preferred as a carrier protein because peptide extensions can be added at either the N-terminus or the C-terminus without affecting folding of the remainder of the molecule. The reason for this is that both the N- and C-termini extend from the folded DHFR molecule in antiparallel ß sheets. Any carrier protein that achieves the desired increase of residence time in the cell could be used instead of DHFR. Thus, plasmid II results in expression of a peptide precursor adjacent to a KEKE motif, both of which are appended to a carrier protein, such as DHFR. The carrier protein element of the fusion protein may become optional if other means of achieving the desired residence time are developed or if sufficient residence time is achieved without the need for a carrier protein.

It is known that direct injection of plasmid DNAs into mammalian muscle can produce cellular immunity. B. Wang et al., 90 Proc. Nat'l Acad. Sci USA 4156 (1993); Z. Xiang et al., 199 Virology 132 (1994); J.

Ulmer et al., 259 Science 1745 (1993). Plasmids I and II would be injected into muscle to produce specific MHC I:peptide complexes for presentation, according to these known methods.

An alternate use of this procedure would be to enhance immunogenicity of tumor specific antigens. For example, malignant melanomas produce a series of specific antigens, called MAGE. G. Nossal, 269(3) Sci. Am. 53, 60 (1993); B. Gaugler et al., 179 J. Exp. Med. 921 (1994). If one could mount significant immunity to MAGE peptides, one might produce an effective therapy for malignant melanoma.

In this second application of the KEKE/activator procedure, lymphocytes are isolated from malignant melanoma patients and then transformed with plasmids I and II. In this case, plasmid II encodes a MAGE antigenic peptide linked to a KEKE motif. These transformed lymphocytes are then reinjected into patients where they would act as cytolytic T lymphocytes for specifically attacking cancer cells. Alternatively, plasmids I and II can be injected intramuscularly, as described.

Other plasmid systems may be used. For example, the pOG series of plasmids (Stratagene) is designed for site-specific integration of foreign DNA into mammalian cells. The pOG plasmids take advantage of FLP recombinase and FLP-Recombination-Targets (FRTs) of *Saccharomyces cerevisiae*, which allow integration of foreign DNA at a specific chromosomal location by site-specific recombination. S. O'Gorman et al., 251 Science 1351 (1991). Mammalian cell lines can be obtained that carry single chromosomally integrated cassettes (pFRTßGAL) that consist of the ß-galactosidase coding sequence, an SV40 early promoter, an SV40 intron, and a polyadenylation signal. The cell line is manipulated so that it constitutively expresses ß-galactosidase activity, however, the ß-

galactosidase gene has been modified to contain an FRT adjacent to the translational start site. The FRT serves as the site of FLP-mediated integration into the chromosome. Integration of foreign DNA results in loss of ß-galactosidase activity and can be screened histochemically, while recombination can be confirmed by hybridization blot analysis. DNA to be integrated into the chromosome is inserted into a targeting vector, such as pOG45, which consists of an FRT and a neomycin resistance cassette in a polylinker-containing phagemid. In the presence of FLP recombinase, site-specific recombination occurs between the chromosomal FRT of pFRTßGAL and the FRT of pOG45, disrupting ß-galactosidase activity and conferring resistance to the drug G418. FLP recombinase is provided by co-transfection with an FLP expression plasmid.

In short, co-expression of activator and appropriate antigenic precursors containing peptides to be presented on MHC I receptors should provide procedures for enhancing cell-mediated immunity against or tolerance to the peptides. By this procedure, KEKE motifs act much like adjuvants do in antibody-mediated responses.

Sequence Listing

(1) GENERAL INFORMATION:

(i) APPLICANT: Martin C. Rechsteiner
Claudio Realini

(ii) TITLE OF INVENTION: Molecular Cloning and Expression of a γ-Interferon Inducible Activator of the Proteasome

(iii) NUMBER OF SEQUENCES: 88

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Thorpe, North & Western
(B) STREET: 9035 South 700 East, Suite 200
(C) CITY: Sandy
(D) STATE: Utah
(E) COUNTRY: USA
(F) ZIP: 84070

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette, 3.5 inch, 720 Kb storage
(B) COMPUTER: AST Advantage NB-SX20
(C) OPERATING SYSTEM: DOS 6.2
(D) SOFTWARE: Word Perfect 5.1

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Alan J. Howarth
(B) REGISTRATION NUMBER: 36,553
(C) REFERENCE/DOCKET NUMBER: T1802

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (801)566-6633
(B) TELEFAX: (801)566-0750

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal fragment

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Homo sapiens*
(G) CELL TYPE: Red blood cells

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Ala Gln Ala Lys Val Asp Val Phe Arg Glu Asp
1 5 10

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal fragment

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Asn Leu Leu Gly Ser Tyr Phe Pro Lys Lys Ile
1 5 10

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal fragment

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Lys Ile Val Val Leu Leu Gln Arg Leu Lys
1 5 10

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal fragment

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Leu Met Thr Ser Leu His Thr Lys
1 5

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal fragment

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Ile Arg Leu Met Val Met Glu Ile Arg Asn Ala Tyr Ala Val Leu Tyr
1 5 10 15

Asp Ile Ile Leu Lys Asn Phe Glu Lys Leu Lys Lys Pro Arg Gly Glu
20 25 30

Thr Lys

(2) INFORMATION FOR SEQ ID N :6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(iii) HYPOTHETICAL: yes

(iv) ANTI-SENSE: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TCGAATTCYT NATGGTNATG GARA 24

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(iii) HYPOTHETICAL: yes

(iv) ANTI-SENSE: yes

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ATAAGCTTTC RTADATCATN CCYTT 25

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(iii) HYPOTHETICAL: yes

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ARYTTYTCRA ARTTYTTNAR GAT 23

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1195 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
AATTCCGTCT CCACCAAAAA AATCGAAAAT TAGTCAGGCG TAGTGGTAGG CACCTGTAAT   60

CCAGCTACTC AGGAGGCTGG TATAGAGAAT CACTGACCCA GGAAGGCGGA GCTGGGTGCG  120

AGCGCCCTAG CTTTCGCTTT CCCTTCGCGG TGCCCACTCC ACTCCTTGTG CGCGCTAGGC  180

CCCGTCCCGG TC ATG GCC ATG CTC AGG GTC CAG CCC GAG GCC CAA GCC AAG  231
             Met Ala Met Leu Arg Val Gln Pro Glu Ala Gln Ala Lys
              1               5                  10

GTG GAT GTG TTT CGT GAA GAC CTC TGT ACC AAG ACA GAG AAC CTG CTC    279
Val Asp Val Phe Arg Glu Asp Leu Cys Thr Lys Thr Glu Asn Leu Leu
    15                  20                  25

GGG AGC TAT TTC CCC AAG AAG ATT TCT GAG CTG GAT GCA TTT TTA AAG    327
Gly Ser Tyr Phe Pro Lys Lys Ile Ser Glu Leu Asp Ala Phe Leu Lys
30                  35                  40                  45
```

```
GAG CCA GCT CTC AAT GAA GCC AAC TTG AGC AAT CTG AAG GCC CCA TTG      375
Glu Pro Ala Leu Asn Glu Ala Asn Leu Ser Asn Leu Lys Ala Pro Leu
                50                  55                  60

GAC ATC CCA GTG CCT GAT CCA GTC AAG GAG AAA GAG AAA GAG GAG CGG      423
Asp Ile Pro Val Pro Asp Pro Val Lys Glu Lys Glu Lys Glu Glu Arg
            65                  70                  75

AAG AAA CAG CAG GAG AAG GAA GAC AAG GAT GAA AAG AAG AAG GGG GAG      471
Lys Lys Gln Gln Glu Lys Glu Asp Lys Asp Glu Lys Lys Lys Gly Glu
        80                  85                  90

GAT GAA GAC AAA GGT CCT CCC TGT GGC CCA GTG AAC TGC AAT GAA AAG      519
Asp Glu Asp Lys Gly Pro Pro Cys Gly Pro Val Asn Cys Asn Glu Lys
    95                  100                 105

ATC GTG GTC CTT CTG CAG CGC TTG AAG CCT GAG ATC AAG GAT GTC ATT      567
Ile Val Val Leu Leu Gln Arg Leu Lys Pro Glu Ile Lys Asp Val Ile
110                 115                 120                 125

GAG CAG CTC AAC CTG GTC ACC ACC TGG TTG CAG CTG CAG ATA CCT CGG      615
Glu Gln Leu Asn Leu Val Thr Thr Trp Leu Gln Leu Gln Ile Pro Arg
                130                 135                 140

ATT GAG GAT GGT AAC AAT TTT GGA GTG GCT GTC CAG GAG AAG GTG TTT      663
Ile Glu Asp Gly Asn Asn Phe Gly Val Ala Val Gln Glu Lys Val Phe
            145                 150                 155

GAG CTG ATG ACC AGC CTC CAC ACC AAG CTA GAA GGC TTC CAC ACT CAA      711
Glu Leu Met Thr Ser Leu His Thr Lys Leu Glu Gly Phe His Thr Gln
        160                 165                 170

ATC TCT AAG TAT TTC TCT GAG CGT GGT GAT GCA GTG ACT AAA GCA GCC      759
Ile Ser Lys Tyr Phe Ser Glu Arg Gly Asp Ala Val Thr Lys Ala Ala
    175                 180                 185

AAG CAG CCC CAT GTG GGT GAT TAT CGG CAG CTG GTG CAC GAG CTG GAT      807
Lys Gln Pro His Val Gly Asp Tyr Arg Gln Leu Val His Glu Leu Asp
190                 195                 200                 205

GAG GCA GAG TAC CGG GAC ATC CGG CTG ATG GTC ATG GAG ATC CGC AAT      855
Glu Ala Glu Tyr Arg Asp Ile Arg Leu Met Val Met Glu Ile Arg Asn
                210                 215                 220

GCT TAT GCT GTG TTA TAT GAC ATC ATC CTG AAG AAC TTC GAG AAG CTC      903
Ala Tyr Ala Val Leu Tyr Asp Ile Ile Leu Lys Asn Phe Glu Lys Leu
            225                 230                 235

AAG AAG CCC AGG GGA GAA ACA AAG GGA ATG ATC TAT TGAGAGCCCT          949
Lys Lys Pro Arg Gly Glu Thr Lys Gly Met Ile Tyr
        240                 245

CTCTCCCATT CTGTGATGAG TAACAGCAGG AGCCTTCCTG CTTTTTACTG GGGACTCCAG 1009

ATTTTCCCCA AAGTTGCTTG TGTTGAGATT TTTCCCTCAC CTTGCCTCTC AGGCTCAATA 1069

AATATAGTTA TACGCTCAGG CTGTGCCCGC AAAGCCTCGG TTGCGTTCCG GTTCCTAGTT 1129

TCCTTCCGGG TGCACGTCGG GGTCGAAGTC AAGGTTGCTC AGGCTCCCAA TAACGACCCG 1189

GGCCGG   1195
```

(2) **INFORMATION FOR SEQ ID NO:10:**

(i) **SEQUENCE CHARACTERISTICS:**
- (A) **LENGTH:** 249 amino acids
- (B) **TYPE:** amino acid
- (D) **TOPOLOGY:** linear

(xi) **SEQUENCE DESCRIPTION:** SEQ ID NO:10:

Met Ala Met Leu Arg Val Gln Pro Glu Ala Gln Ala Lys Val Asp Val
1 5 10 15

Phe Arg Glu Asp Leu Cys Thr Lys Thr Glu Asn Leu Leu Gly Ser Tyr
20 25 30

Phe Pro Lys Lys Ile Ser Glu Leu Asp Ala Phe Leu Lys Glu Pro Ala
35 40 45

Leu Asn Glu Ala Asn Leu Ser Asn Leu Lys Ala Pro Leu Asp Ile Pro
50 55 60

Val Pro Asp Pro Val Lys Glu Lys Glu Lys Glu Glu Arg Lys Lys Gln
65 70 75 80

Gln Glu Lys Glu Asp Lys Asp Glu Lys Lys Lys Gly Glu Asp Glu Asp
85 90 95

Lys Gly Pro Pro Cys Gly Pro Val Asn Cys Asn Glu Lys Ile Val Val
100 105 110

Leu Leu Gln Arg Leu Lys Pro Glu Ile Lys Asp Val Ile Glu Gln Leu
115 120 125

Asn Leu Val Thr Thr Trp Leu Gln Leu Gln Ile Pro Arg Ile Glu Asp
130 135 140

Gly Asn Asn Phe Gly Val Ala Val Gln Glu Lys Val Phe Glu Leu Met
145 150 155 160

Thr Ser Leu His Thr Lys Leu Glu Gly Phe His Thr Gln Ile Ser Lys
165 170 175

Tyr Phe Ser Glu Arg Gly Asp Ala Val Thr Lys Ala Ala Lys Gln Pro
180 185 190

His Val Gly Asp Tyr Arg Gln Leu Val His Glu Leu Asp Glu Ala Glu
195 200 205

Tyr Arg Asp Ile Arg Leu Met Val Met Glu Ile Arg Asn Ala Tyr Ala
210 215 220

Val Leu Tyr Asp Ile Ile Leu Lys Asn Phe Glu Lys Leu Lys Lys Pro
225 230 235 240

Arg Gly Glu Thr Lys Gly Met Ile Tyr
245

(2) **INFORMATION FOR SEQ ID NO:11:**

(i) **SEQUENCE CHARACTERISTICS:**
(A) **LENGTH:** 28 amino acids
(B) **TYPE:** amino acid
(D) **TOPOLOGY:** linear

(xi) **SEQUENCE DESCRIPTION:** SEQ ID NO:11:

Lys Glu Lys Glu Lys Glu Glu Arg Lys Lys Gln Gln Glu Lys Glu Asp
1 5 10 15

Lys Asp Glu Lys Lys Lys Gly Glu Asp Glu Asp Lys
20 25

(2) **INFORMATION FOR SEQ ID NO:12:**

(i) **SEQUENCE CHARACTERISTICS:**
(A) **LENGTH:** 24 amino acids
(B) **TYPE:** amino acid
(D) **TOPOLOGY:** linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Lys Lys His Glu Glu Glu Glu Ala Lys Ala Glu Arg Glu Lys Lys Glu
1 5 10 15

Lys Glu Gln Arg Glu Lys Asp Lys
20

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Glu Lys Lys Glu Gly Gln Glu Lys Glu Glu Ser Lys Lys Asp Arg Lys
1 5 10 15

Glu Asp Lys Glu Lys Asp Lys Asp Lys Glu Lys Ser Asp Val Lys Lys
20 25 30

Glu Lys Lys
35

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Lys Ile Ile Glu Lys Glu Lys Glu Glu Glu Leu Glu Lys Lys Lys Gln
1 5 10 15

Lys

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Glu Glu Lys Glu Asp Lys Glu Glu Glu Lys Glu Lys Glu Glu Lys Glu
1 5 10 15

Ser Glu Asp Lys
20

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Glu Lys Leu Ala Ala Gln Arg Lys Ala Glu Ala Glu Lys Lys Glu Glu
1 5 10 15

Lys Lys Asp Thr Glu
20

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Glu Glu Glu Lys Lys Ala Ala Glu Glu Arg Ala Lys Ala
1 5 10

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This fluorogenic tetrapeptide is succinylated at the N-terminus and contains 7-amido-4-methylcoumarin at the C-terminus.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Leu Leu Val Tyr
1

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This fluorogenic octapeptide is succinylated at the N-terminus and contains 7-amido-4-methylcoumarin at the C-terminus.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Arg Pro Phe His Leu Leu Val Tyr
1 5

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This fluorogenic pentapeptide is succinylated at the N-terminus and contains 7-amido-4-methylcoumarin at the C-terminus.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Gly Pro Leu Gly Pro
1 5

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) **FEATURE:**
(D) **OTHER INFORMATION:** This peptide is fused at its N-terminus to the C-terminus of ubiquitin (Ub).

(xi) **SEQUENCE DESCRIPTION:** SEQ ID NO:21:

Pro Val Lys Glu Lys Glu Lys Glu Glu Arg Lys Lys Gln Gln Glu Lys
1 5 10 15

Glu Asp Lys Asp Glu Lys Lys Lys Gly Glu Asp
20 25

(2) **INFORMATION FOR SEQ ID NO:22:**

(i) **SEQUENCE CHARACTERISTICS:**
(A) **LENGTH:** 9 amino acids
(B) **TYPE:** amino acid
(D) **TOPOLOGY:** linear

(ix) **FEATURE:**
(D) **OTHER INFORMATION:** This peptide borne by JAK 1 kinase is presented on MHC class I receptors.

(xi) **SEQUENCE DESCRIPTION:** SEQ ID NO:22:

Ser Phe Phe Pro Glu Ile Thr His Ile
1 5

(2) **INFORMATION FOR SEQ ID NO:23:**

(i) **SEQUENCE CHARACTERISTICS:**
(A) **LENGTH:** 42 amino acids
(B) **TYPE:** amino acid
(D) **TOPOLOGY:** linear

(ix) **FEATURE:**
(D) **OTHER INFORMATION:** Residues 1-29 constitute the KEKE sequence, and residues 34-42 are the presented peptide.

(xi) **SEQUENCE DESCRIPTION:** SEQ ID NO:23:

Lys Glu Lys Glu Lys Asn Lys Leu Lys Arg Lys Lys Leu Glu Asn Lys
1 5 10 15

Asp Lys Lys Asp Glu Glu Lys Asn Lys Ile Arg Glu Glu Trp Asn Asn
20 25 30

Phe Ser Phe Phe Pro Glu Ile Thr His Ile
35 40

(2) **INFORMATION FOR SEQ ID NO:24:**

(i) **SEQUENCE CHARACTERISTICS:**
(A) **LENGTH:** 33 amino acids
(B) **TYPE:** amino acid
(D) **TOPOLOGY:** linear

(xi) **SEQUENCE DESCRIPTION:** SEQ ID NO:24:

Lys Glu Glu Glu Glu Glu Lys Glu Glu Glu Lys Asp Lys Gly Asp Glu
1 5 10 15

Glu Glu Glu Gly Glu Glu Lys Leu Glu Glu Lys Gln Lys Ser Asp Ala
20 25 30

Glu

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Lys Glu Lys Glu Lys Asn Lys Leu Lys Arg Lys Lys Leu Glu Asn Lys
1 5 10 15

Asp Lys Lys Asp Glu Glu Lys Asn Lys Ile Arg Glu Glu
20 25

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:

(D) OTHER INFORMATION: Residues 1-13 constitute the KEKE sequence, and residues 24-33 are the presented peptide.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Glu Glu Glu Lys Lys Lys Met Glu Glu Ser Lys Ala Lys Phe Glu Asn
1 5 10 15

Leu Cys Lys Leu Met Lys Glu Ile Leu Asp Lys Lys Val Glu Lys Val
20 25 30

Thr

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Leu Ser Pro Phe Pro Phe Asp Leu
1 5

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Thr Gln His Asn Arg Ala Leu Asp Leu
1 5

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: Residues 1-17 constitute the KEKE sequence, and residues 25-33 are the presented peptide.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Glu Lys Leu Asp Ile Lys Leu Asp Ser Glu Asp Lys Asp Lys Glu Gly
1 5 10 15

Lys Pro Leu Leu Lys Ala Val Met Arg Arg Trp Leu Pro Ala Gly Asp
20 25 30

Ala

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: Residues 1-13 constitute the KEKE sequence, and residues 28-36 are the presented peptide.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Lys Lys Glu Lys Ala Arg Val Ile Thr Glu Glu Glu Lys Asn Phe Lys
1 5 10 15

Ala Phe Ala Ser Leu Arg Met Ala Arg Ala Asn Ala Arg Leu Phe Gly
20 25 30

Ile Arg Ala Lys
35

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: Residues 1-13 constitute the KEKE sequence, and residues 26-34 are the presented peptide.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Glu Thr Glu Asp Asn Lys Glu Lys Lys Ser Ala Lys Asp Ala Leu Leu
1 5 10 15

Leu Trp Cys Gln Met Lys Thr Ala Gly Tyr Pro Asn Val Asn Ile His
20 25 30

Asn Phe

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Ile Tyr Ala Thr Val Ala Gly Ser
1 5

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Ser Ser Glu Gln Thr Phe Met Tyr Tyr
1 5

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Arg Gly Tyr Val Tyr Gln Gly Leu
1 5

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 72 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: Residues 4-16 constitute the KEKE sequence. This peptide is linked at its C-terminus to SEQ ID NO:36 through a sequence of variable length.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Lys Pro Ala Glu Lys Lys Asp Asp Leu Lys Glu Glu Lys Lys Asp Asp
1 5 10 15

Leu Pro Lys Glu Glu Lys Lys Asp Asp Leu Pro Lys Glu Glu Lys Lys
20 25 30

Asp Asp Pro Pro Lys Glu Glu Lys Lys Asp Asp Leu Pro Lys Glu Glu
35 40 45

Lys Lys Asp Ala Pro Lys Asp Gly Asn Lys Asp Ala Por Lys Glu Glu
50 55 60

Lys Lys Ala Asp Pro Pro Lys Glu
65 70

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:

(D) OTHER INF RMATION: This peptide constitutes a presented peptide and is linked at its N-terminus to SEQ ID NO:35 through a sequence of variable length.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Glu Asp Ser Tyr Val Pro Ser Ala Glu Gln Ile Leu Glu Phe Val Lys
1 5 10 15

Gln Met

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This peptide constitutes a KEKE sequence, and is linked at its C-terminus to SEQ ID NO: 38 by a proline-rich region.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Glu Gly Lys Lys Asn Glu Lys Lys Asn Glu Lys Ile Glu Arg Asn Asn
1 5 10 15

Lys

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This peptide is linked at its N-terminus to SEQ ID NO:37 through a proline-rich region. Residues 12-20 constitute the presented peptide.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Asn Asn Asp Asp Ser Tyr Ile Pro Ser Ala Glu Lys Ile Leu Glu Phe
1 5 10 15

Val Lys Gln Ile
20

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This peptide constitutes a KEKE sequence, and is linked to SEQ ID NO:40 through a sequence of variable number.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Lys Glu Gly Ala Asp Lys Glu Lys Lys Lys Glu Lys Gly Lys Glu Lys
1 5 10 15

Glu Glu Glu

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This peptide is linked to SEQ ID NO:39 through a sequence of variable number. Residues 12-19 constitute a presented peptide.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Asn Glu Lys Val Val Asn Asp Tyr Leu Leu His Lys Ile Arg Ser Ser
1 5 10 15

Val Thr Thr

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This peptide constitutes a presented peptide, and is linked at its C-terminus to SEQ ID NO:42 through a sequence of 17 amino acids.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Glu Tyr Leu Asn Lys Ile Gln Asn Ser Leu Ser Thr Glu Trp Ser Pro
1 5 10 15

Cys Ser Val Thr
20

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This peptide constitutes a KEKE-like sequence and is linked at its N-terminus to SEQ ID NO:41 through a sequence of 17 amino acids.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

Lys Asp Glu Leu Asp Tyr Ala Asn Asp Ile Glu Lys Lys Ile Cys Lys
1 5 10 15

Met Glu Lys

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) T P LOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This peptide constitutes a KEKE sequence, and is linked at its C-terminus to SEQ ID NO:44 through a glutamine-rich region.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Lys Asp Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys
1 5 10 15

Ser Lys Lys Lys
20

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This peptide constitutes a presented peptide, and is linked at its N-terminus to SEQ ID NO:43 through a glutamine-rich region.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Arg Trp Ile Ile Leu Gly Leu Asn Lys
1 5

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

Arg Arg Ser Lys Glu Ile Thr Val Arg
1 5

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

Gly Tyr Lys Asp Gly Asn Glu Tyr Ile
1 5

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This peptide constitutes a KEKE-like sequence, and is linked at its C-terminus to SEQ ID NO:48 through a sequence of 60 amino acids.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Arg Ala Glu Glu Glu Asp Lys Lys Glu Asp
1 5 10

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This peptide is linked at its N-terminus to SEQ ID NO:47 through a sequence of 60 amino acids. Residues 3-13 constitute the presented peptide.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Glu Asn Thr Val Phe Asp Ala Lys Arg Leu Ile Gly Arg
1 5 10

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

Arg Pro Gln Ala Ser Gly Val Tyr Met
1 5

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

Ser Ser Ile Glu Phe Ala Arg Leu
1 5

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This peptide constitutes a KEKE-like sequence, and is linked at its C-terminus to SEQ ID NO:52 through a sequence of 34 amino acids.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

Glu Glu Glu Gly Lys Gly Lys Asp Ala Ser Gly Asn Lys Val Lys Ala
1 5 10 15

Glu

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This peptide is linked at its N-terminus to SEQ ID NO:51 through a sequence of 34 amino acids. Residues 5-13 constitute the presented peptide.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

Gly Val Asn Leu Pro Gln Lys Ala Gly Gly Phe Leu Met
1 5 10

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

Val Pro Lys Leu Lys Val Cys Ala Leu
1 5

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This peptide, constituting a KEKE sequence, is linked at its C-terminus to SEQ ID NO:55 through a sequence of 62 amino acids.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

Lys Glu Ile Cys Glu Lys Asn Asp Glu Cys Glu Ser Ser Lys Glu
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This peptide is linked at its N-terminus to SEQ ID NO:54 through a sequence of 62 amino acids. Residues 8-16 constitute the presented peptide.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

Asn Ile Arg Thr Leu Ile Gln Ile Leu Lys Gln Lys Ile Ala Asp Leu
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

Val Ser Asp Ile Val Gly Pro Asp Gly Leu Val Tyr
1 5 10

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This peptide, constituting a KEKE-like sequence, is linked at its C-terminus to SEQ ID NO:58 through a sequence of 15 amino acids.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

Glu Met Lys Lys Arg Glu Ser Lys Phe Ile Lys Asp Ala Asp Glu Glu
1 5 10 15
Lys

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This peptide, constituting a KEKE-like sequence, is linked at its N-terminus to SEQ ID NO:57 through a sequence of 15 amino acids and at its C-terminus to SEQ ID NO:59 through a sequence of 24 amino acids.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

Glu Lys Asp Ala Glu Leu Glu Lys Leu Arg Asn Glu
1 5 10

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This peptide is linked at its N-terminus to SEQ ID NO:58 through a sequence of 24 amino acids. Residues 5-13 are the presented peptide.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

Lys Val Lys Leu Glu Leu Lys Val Lys Asn Leu Glu Leu
1 5 10

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

Gly Ile Leu Gly Phe Val Phe Thr Leu
1 5

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

Ser Ala Ile Asn Asn Tyr Ala Gln Lys Leu
1 5 10

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

Cys Lys Gly Val Asn Lys Glu Tyr Leu
1 5

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

Ser Ile Ile Asn Phe Glu Lys Leu
1 5

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This peptide constitutes a presented peptide and is linked at its C-terminus to SEQ ID NO:65 through a sequence of 13 amino acids.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

Arg Arg Ile Lys Glu Ile Val Lys Lys
1 5

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This peptide, constituting a KEKE sequence, is linked at its N-terminus to SEQ ID NO:64 through a sequence of 13 amino acids.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

Glu Lys Glu Arg Asp Lys Glu Val Ser Asp Asp Glu Ala Glu Lys Glu
1 5 10 15

Asp Lys Glu Glu Glu Lys Glu Lys Glu Glu Lys Glu Ser Glu Asp Lys
20 25 30

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: Residues 1-9 constitute the presented peptide, and residues 9-23 constitute the KEKE sequence.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

Ala Arg Leu Phe Gly Ile Arg Ala Lys Arg Ala Lys Glu Ala Ala Glu
1 5 10 15

Gln Asp Val Glu Lys Lys Lys
20

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: Residues 1-9 constitute the presented peptide, and residues 10-22 constitute the KEKE-like sequence.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

Lys Tyr Gln Ala Val Thr Thr Thr Leu Glu Glu Lys Arg Lyr Glu Lys
1 5 10 15

Ala Lys Ile His Tyr Arg
20

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

Phe Ala Pro Gly Asn Tyr Pro Ala Leu
1 5

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

Met Ile Glu Pro Arg Thr Leu Gln Tyr
1 5

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

Leu Leu Phe Gly Tyr Pro Val Tyr Val
1 5

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

Pro Lys Tyr Val Lys Gln Asn Thr Leu Lys Leu Ala
1 5 10

(2) INFORMATION FOR SEQ ID NO:72:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

Thr Leu Trp Val Asp Pro Tyr Glu Val
1 5

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

Glu Ala Asp Pro Thr Gly His Ser Tyr Val
1 5 10

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

Leu Tyr Gln Asn Val Gly Thr Tyr Val
1 5

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

Arg Arg Tyr Pro Asp Ala Val Tyr Leu
1 5

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

His Pro Lys Tyr Lys Thr Glu Leu
1 5

(2) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

Glu Pro Lys Tyr Lys Thr Gln Leu
1 5

(2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

Ala Ser Asn Glu Asn Met Glu Thr Met
1 5

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

Ser Gly Pro Ser Asn Thr Pro Pro Glu Ile
1 5 10

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

Ser Thr Gly Asn Leu Ile Ala Pro Glu Tyr Gly Phe Lys Ile Ser
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

Arg Tyr Leu Lys Asn Gly Lys Glu Thr Leu Gln Arg Ala
1 5 10

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

Gly Arg Ile Asp Lys Pro Ile Leu
1 5

(2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

Phe Arg Tyr Asn Gly Leu Ile His Arg
1 5

(2) INFORMATION FOR SEQ ID NO:84:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

Gly Arg Ala Phe Val Thr Ile Gly Lys
1 5

(2) INFORMATION FOR SEQ ID NO:85:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

Thr Tyr Gln Arg Thr Arg Ala Leu Val
1 5

(2) INFORMATION FOR SEQ ID NO:86:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 70 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

TTTCTCCCCT GGGCTTCTTG AGCTTCTCGA AGTTCTTCAG GATGATGTCA 50

TATAACACAG CATAAGCATT 70

(2) INFORMATION FOR SEQ ID NO:87:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This peptide constitutes the presented peptide and is linked at its C-terminus to SEQ ID NO: 88 through a sequence of about 150 amino acids.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

Gln Gly Ile Asn Asn Leu Asp Asn Leu
1 5

(2) INFORMATION FOR SEQ ID NO:88:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: This peptide constitutes a KEKE-like sequence and is linked at its N-terminus to SEQ ID NO: 87 through a sequence of about 150 amino acids.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

Glu Asp Ser Gln Glu Asn Ala Asp Lys Asn Glu Asp Gly Gly Glu Lys
1 5 10 15

Claims

We claim:

1. A purified polynucleotide having a nucleotide sequence that encodes a proteasome activator, wherein said proteasome activator is of human origin, has an M_r of about 29,000, and is derived from a hexameric activator complex containing 29 kD and 31 kD subunits.

2. The polynucleotide of Claim 1 wherein the nucleotide sequence comprises SEQ ID NO:9.

3. The polynucleotide of Claim 2 wherein said polynucleotide encodes a protein comprising an amino acid sequence identified as SEQ ID NO:10.

4. A protein capable of activating proteasomes *in vitro*, wherein said protein has an M_r of about 29,000 and is expressed from a transformable polynucleotide having a nucleotide sequence encoding a 29 kD subunit of a hexameric human activator complex comprised of 29 kD and 31 kD subunits.

5. The protein of Claim 4 wherein said nucleotide sequence comprises SEQ ID NO:9.

6. The protein of Claim 5 wherein said protein has an amino acid sequence identified as SEQ ID NO:10.

7. A method of activating proteasomes comprising the step of contacting the proteasomes with a protein expressed from a transformable polynucleotide having a nucleotide sequence encoding a 29 kD subunit of a hexameric human activator complex comprised of 29 kD and 31 kD subunits, wherein said contacting occurs under conditions suitable for binding of said protein to said proteasomes.

8. The method of Claim 7 wherein said nucleotide sequence comprises SEQ ID NO:9.

9. The method of Claim 8 wherein said protein has an amino acid sequence comprising SEQ ID NO:10.

10. A method for inducing synthesis of a proteasome activator in cultured human cells, wherein said activator comprises a 29 kD subunit of a hexameric human activator complex comprised of 29 kD and 31 kD subunits, comprising the step of contacting said cells with an effective amount of γ-interferon.

11. The method of Claim 10 wherein said activator has an amino acid sequence identified as SEQ ID NO:10.

12. The method of Claim 11 wherein said amino acid sequence is encoded by SEQ ID NO:9.

13. A method for eliciting cell-mediated immunity or tolerance to a selected immunogenic peptide in a warm-blooded animal comprising

(a) providing at least one plasmid encoding a proteasome activator and a precursor peptide containing the immunogenic peptide and at least one KEKE motif-containing peptide, wherein said KEKE motif-containing peptide is positioned adjacent to said immunogenic peptide in said precursor peptide and said plasmid expresses both said proteasome activator and said precursor peptide upon insertion into an appropriate host cell; and

(b) injecting said plasmid into an appropriate site in said warm-blooded animal.

14. The method of claim 13 wherein said proteasome activator has the amino acid sequence identified as SEQ ID NO:10.

15. The method of claim 14 wherein said immunogenic peptide is selected from the group consisting of presented peptides of pathogens and tumor antigens.

16. The method of claim 15 wherein said immunogenic peptide is a presented peptide of a pathogen.

17. The method of claim 15 wherein said immunogenic peptide is a tumor antigen.

18. The method of claim 15 wherein said proteasome activator and said precursor peptide are encoded on and expressed from separate plasmids.

19. The method of claim 15 wherein said plasmid becomes integrated into a chromosome of said warm-blooded animal.

20. The method of claim 15 wherein said KEKE motif-containing peptide is SEQ ID NO:11.

21. The method of claim 15 wherein said precursor peptide further comprises a carrier protein.

22. The method of claim 21 wherein said carrier protein is dihydrofolate reductase.

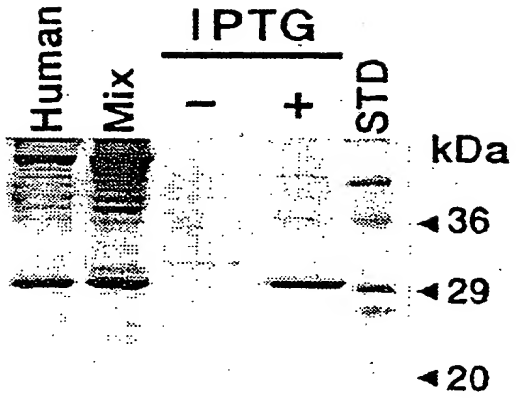

Fig. 1A

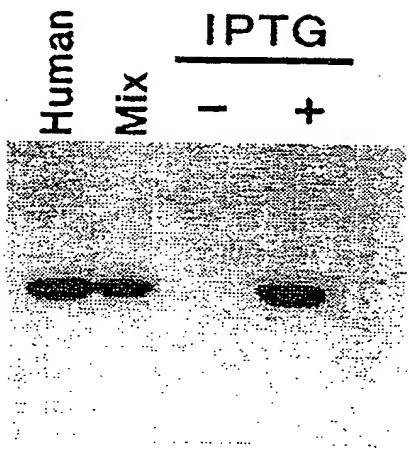

Fig. 1B

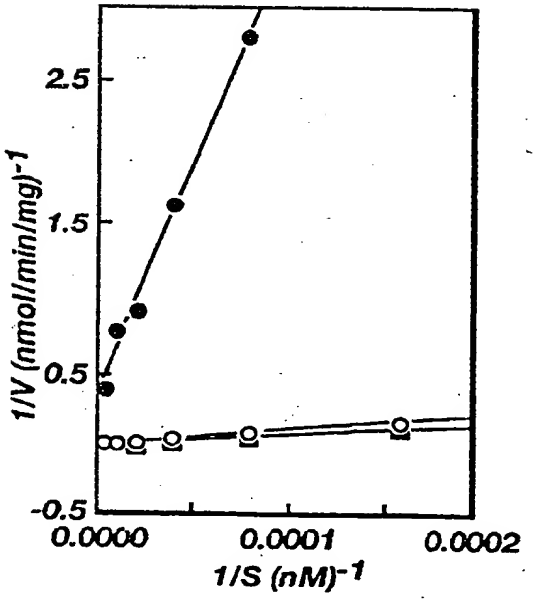

Fig. 3A

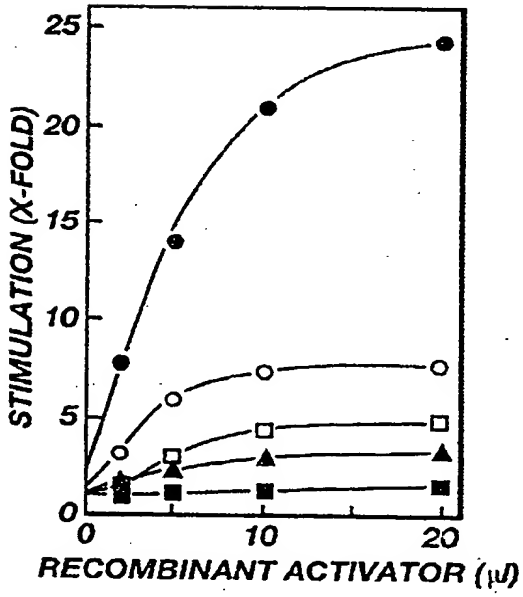

Fig. 3B

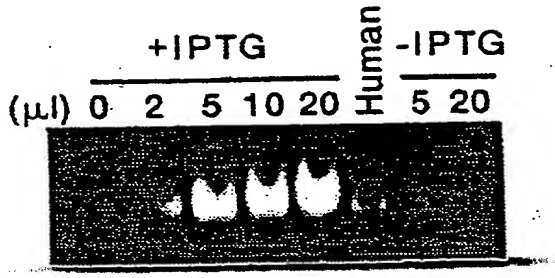

Fig. 2A

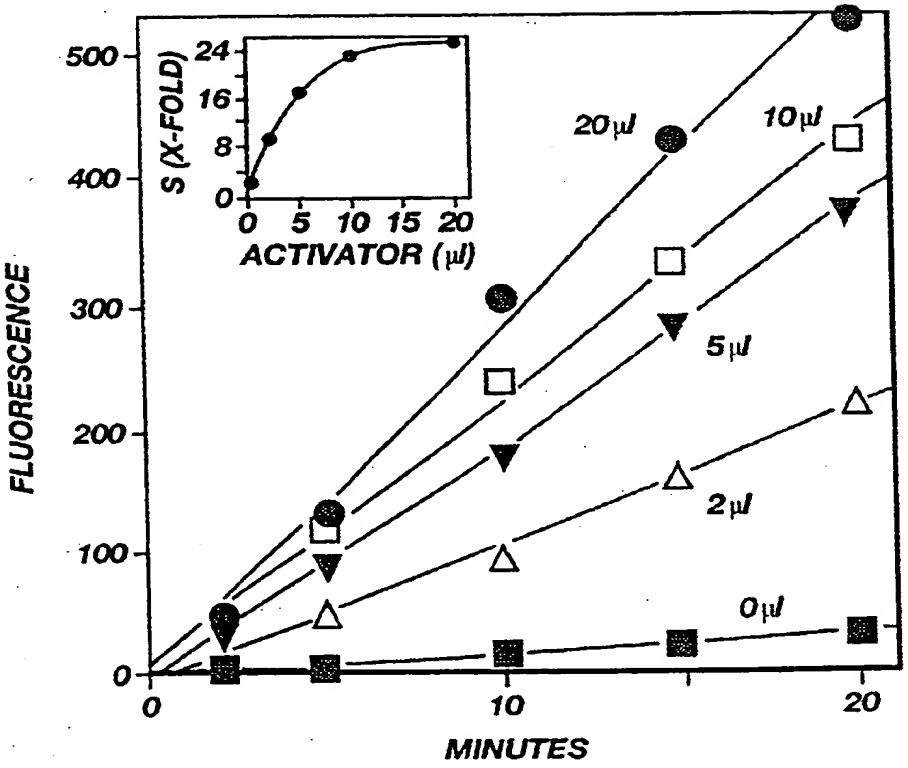

Fig. 2B

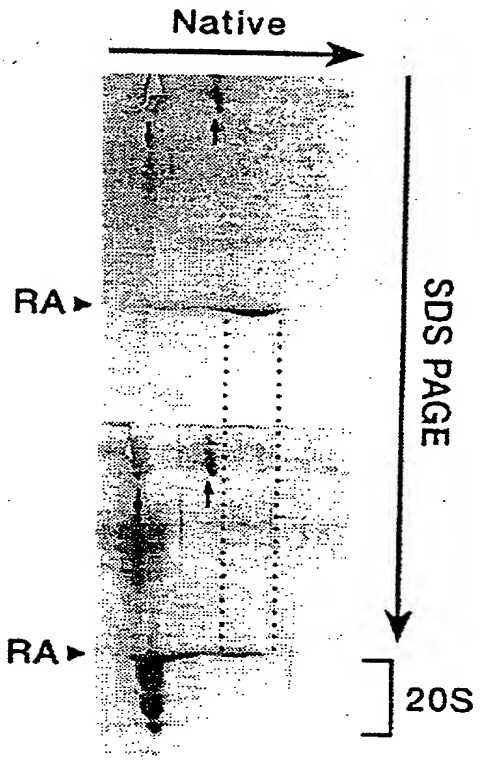

Fig. 2C

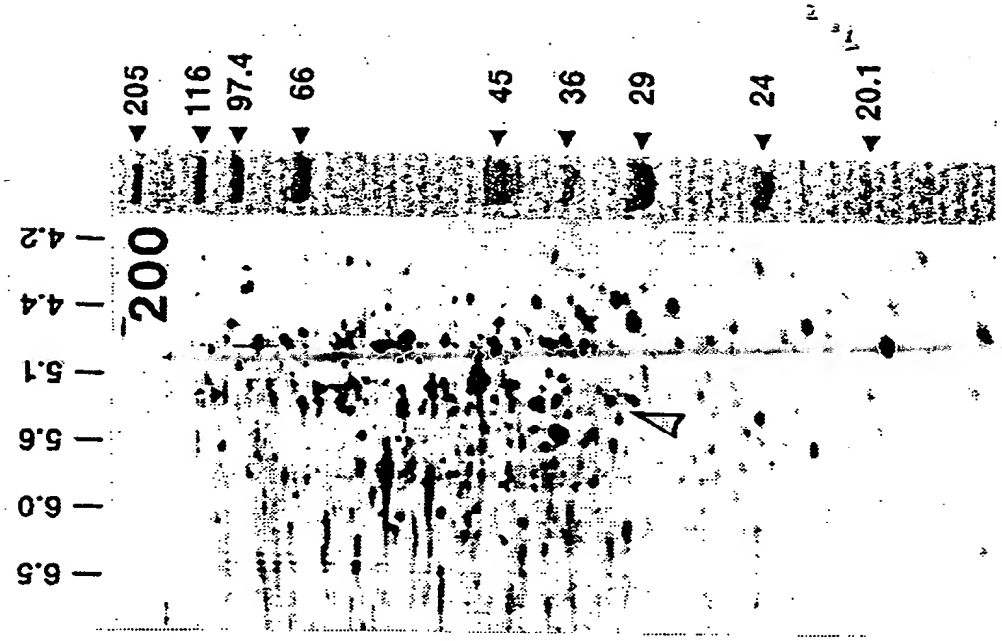

Fig. 4C

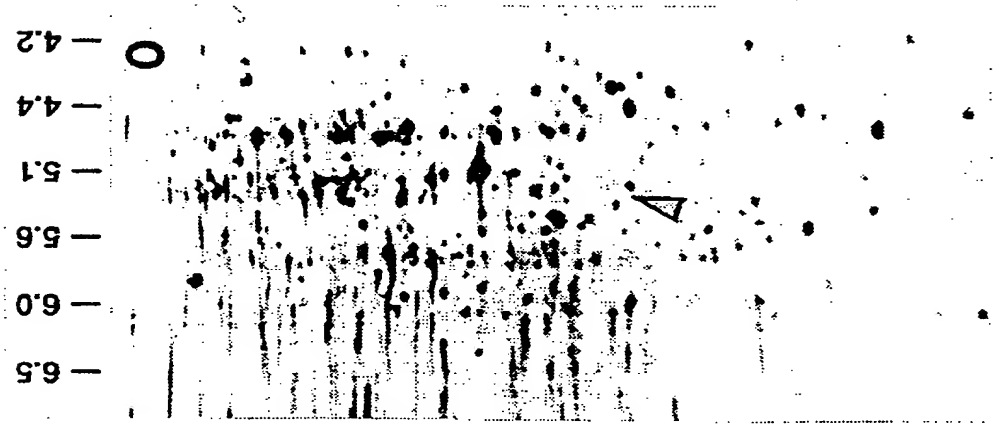

Fig. 4B

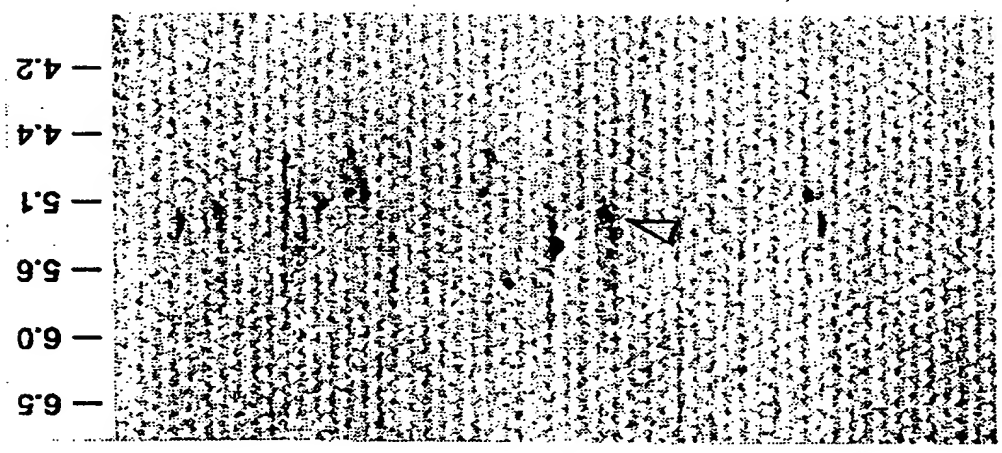

Fig. 4A

```
                1               5                  10                  15                  20                  25                  30                  35
SEQ ID NO:11    Lys Glu Lys Glu Lys Glu Glu Arg Lys Lys Gln Gln Glu Lys Glu Asp Lys Asp Glu Lys Lys Lys Glu Asp Lys Glu Lys Gly Glu Asp Lys Ser Asp Val Lys Lys Glu Lys Lys
SEQ ID NO:12    Lys Lys His Glu Glu Glu Ala Lys Ala Glu Arg Glu Lys Lys Glu Lys Glu Gln Arg Glu Lys Asp
SEQ ID NO:13    Glu Lys Lys Gly Gln Glu Lys Glu Ser Lys Lys Asp Arg Lys Glu Asp Lys Asp Lys Glu Asp Lys
SEQ ID NO:14    Lys Ile Ile Glu Lys Glu Glu Leu Glu Lys Lys Lys Gln Lys Glu Lys Asp Lys Glu Lys Ser Asp
SEQ ID NO:15    Glu Glu Lys Glu Asp Lys Glu Glu Lys Glu Lys Glu Glu Lys Glu Ser Glu Lys Asp Lys Glu Lys
SEQ ID NO:16    Glu Lys Leu Ala Ala Gln Arg Lys Ala Glu Ala Glu Lys Lys Glu Lys Lys Glu Lys Glu Lys Glu
SEQ ID NO:24    Lys Glu Glu Glu Glu Lys Glu Glu Lys Asp Lys Gly Asp Glu Glu Glu Lys Asp Lys Glu Glu Lys
SEQ ID NO:25    Lys Glu Lys Glu Lys Asn Lys Leu Lys Arg Lys Lys Leu Glu Asn Lys Asp Lys Asp Ala Glu Glu
```

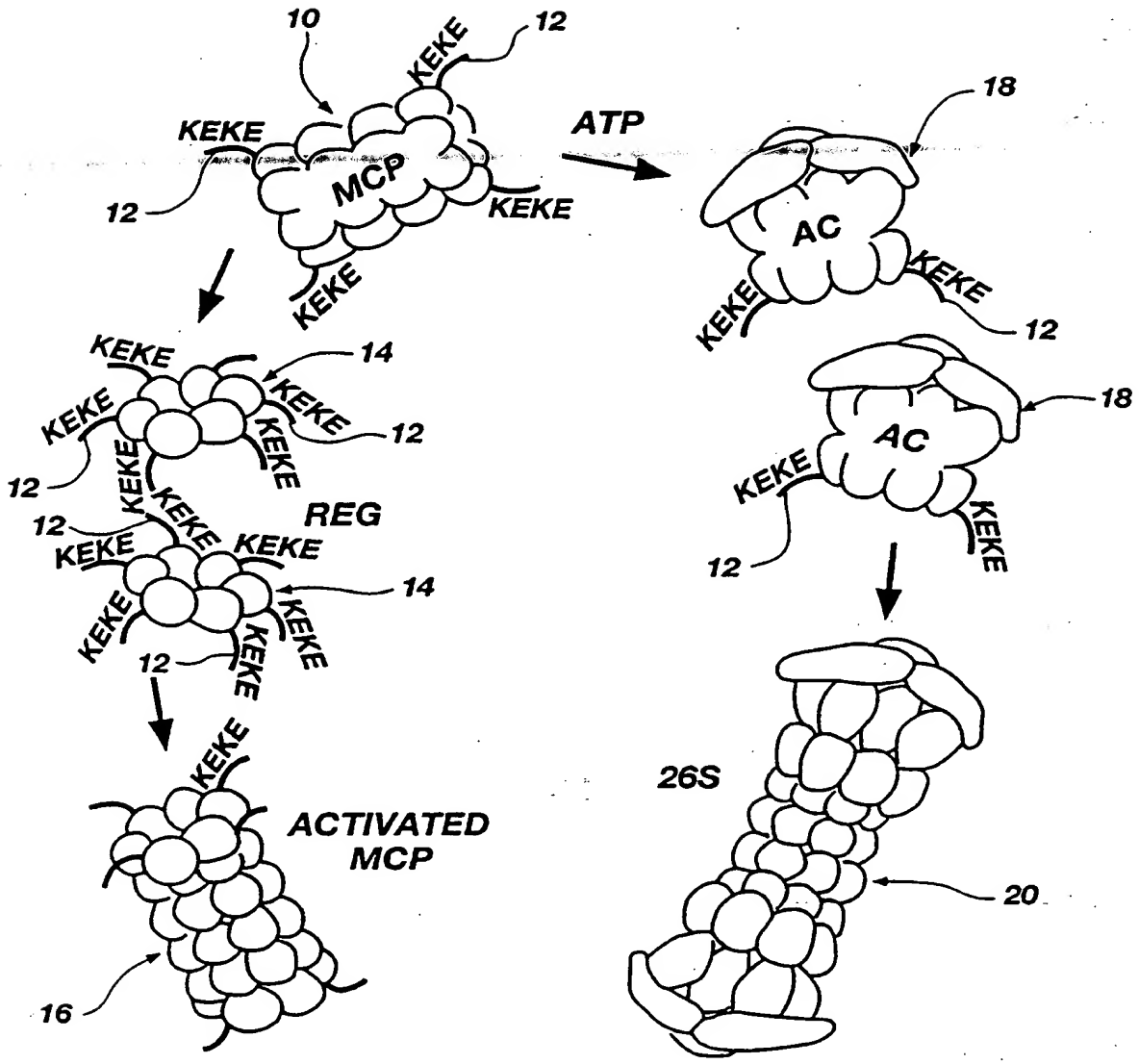

Fig. 6

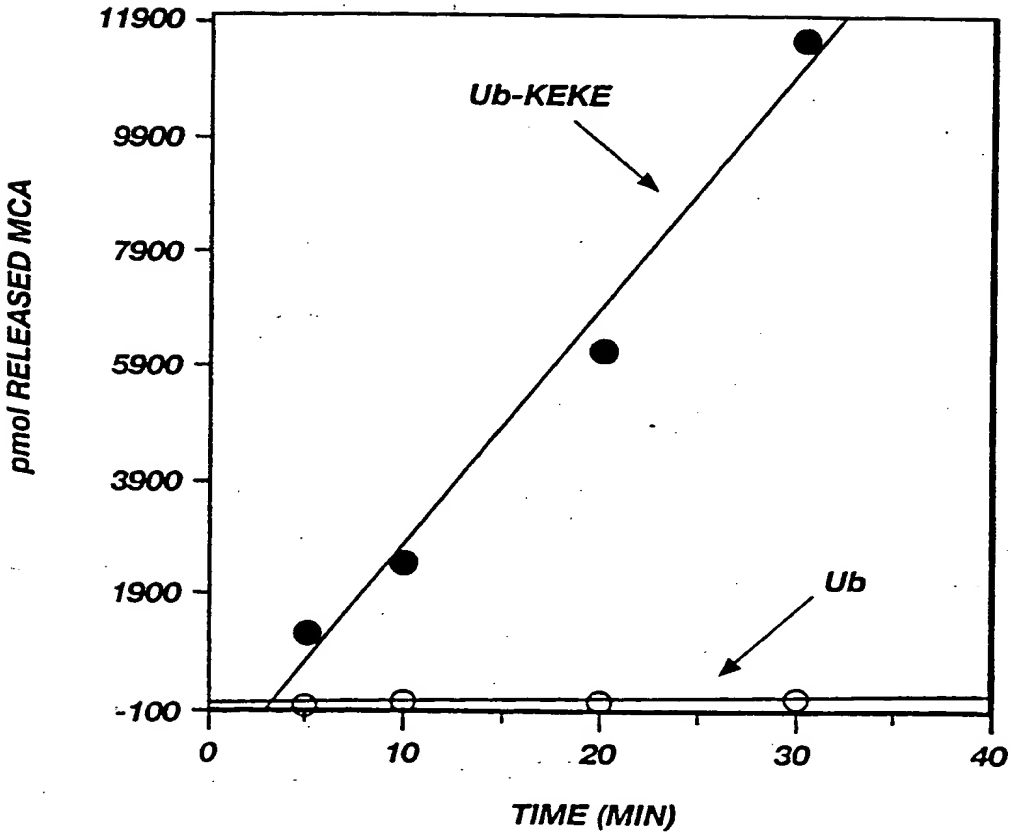

Fig. 7

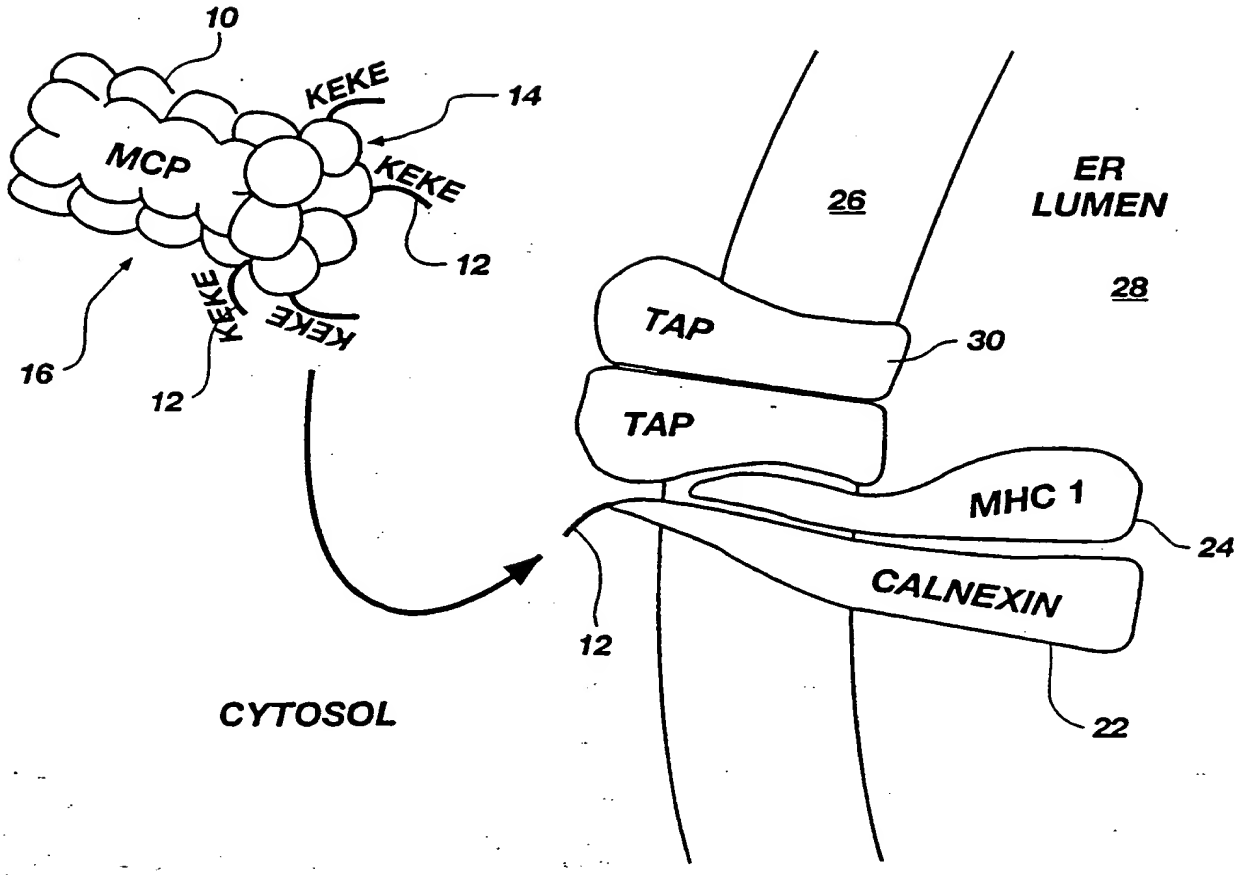

Fig. 8

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/03591

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :C12N 15/12, 15/09, 9/64; C07K 13/00; A61K 39/385, 39/39, 48/00
US CL :536/23.5; 530/350, 351; 435/226, 60.2; 424/192.1, 278.1; 514/44
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.5; 530/350, 351; 435/226, 60.2; 424/192.1, 278.1; 514/44

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, CAS, MEDLINE, BIOSIS , LIFESCI, BIOTECHDS, WPIDS, EMBASE
search terms: proteasome# or macropain or muticatalytic protease# or 20S protease# or 26S protease#, activat?, KEKE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y --- A	Journal of Biological Chemistry, Volume 267, No. 31, issued 05 November 1992, W. Dubiel et al., "Purification of an 11S Regulator of the Multicatalytic Protease", pages 22369-22377, see entire document.	4-9 ----- 1-3 ----- 13-22
Y --- A	Journal of Biological Chemistry, Volume 267, No. 15, issued 25 May 1992, M. Chu-Ping et al., "Identification, Purification, and Characterization of a Protein Activator (PA28) of the 20 S Proteasome (Macropain)", pages 10515-10523, see entire document.	1-9 ----- 13-22

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
16 SEPTEMBER 1994	24 OCT 1994
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer REBECCA PROUTY Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/03591

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y --- A	Biochemical and Biophysical Research Communications, Volume 178, No. 1, issued 15 July 1991, M. Yukawa et al., "Proteasome and Its Novel Endogenous Activator in Human Platelets", pages 256-262, see particularly pages 259-262.	4-9 --- 1-3 --- 13-22
X --- Y --- A	Journal of Biochemistry, Volume 114, No. 3, issued September 1993, M. Yukawa et al., "Purification and Characterization of Endogenous Protein Activator of Human Platelet Proteasome", pages 317-323, see entire document.	4-9 --- 1-3 --- 13-22
Y --- A	Biological Chemistry Hoppe-Seyler, Volume 374, No. 9, issued September 1993, L. Kuehn et al., "Purification and Some Properties of an Endogenous Activator of the Multicatalytic Proteinase (Proteasome) From Rabbit Red Blood Cells", page 710, see entire document.	4-9 --- 1-3 --- 13-22
X	European Journal of Biochemistry, Volume 218, issued December 1993, B. Honore et al., "Interferon-γ Up Regulates a Unique set of Proteins in Human Keratinocytes. Molecular Cloning and Expression of the cDNA Encoding the RGD-Sequence-Containing Protein IGUP I-5111", pages 421-430, see entire document.	1-6, 10-12

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/03591

Box I Observations where certain claims were found unsearchable (Continuation of item 1 f first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark n Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/03591

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

Group I, claim(s)1-3 and 13-22, drawn to DNA encoding proteasome activating factor and method of use of the DNA.
Group II, claim(s) 4-9, drawn to proteasome activating factor and method of use of the protein.
Group III, claim(s) 10-12, drawn to a method of inducing the synthesis of proteasome activating factor.

The inventions listed as Groups I, II and III do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:
The DNA of Group I and protein of Group II are chemically distinct compounds composed of different constituents.
The methods of Group III do not utilize the compounds of either Group I or Group II.